Proceedings

Thirteenth Annual

Ground Target Modeling and Validation Conference

August 2002

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Hosted by
Signature Research, Inc.

Sponsored by
U.S. Army TARDEC

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT DATE	3. DATES COVERED *(From - To)*
24-02-2003	Conference Proceedings February 2003	August 2002

4. TITLE AND SUBTITLE

Proceedings of the Thirteenth Annual Ground Target Modeling and Validation Conference

(Unclassified)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Various

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Signature Research Inc.
P.O. Box 346
Calumet MI 49913

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army-Tank Automotive Command Research, Development and Engineering Center
ATTN: AMSTA-TR-R
Warren MI 48397-5000

10. SPONSOR/MONITOR'S ACRONYM(S)
US Army TARDEC

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT

Unlimited (1)

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This group of papers represents the material presented at the 2002 Conference on Ground Target Modeling and Validation.
The conference serves as a workshop brining members of the defense research community together to discuss current developments in ground target modeling and validation. It is open to DoD employees, defense industry contractors, academia, and foreign nations.
The focus of the conference is on scientific and technological advancements relating to the tools and techniques of modeling, and the simulation of ground target signatures and background scenes for infrared or near-infrared, radar/millimeter wave, acoustic, seismic, photometric, ultraviolet, and magnetic testing.

15. SUBJECT TERMS
Signature, Countermeasures, Military Targets, Modeling Validation, Testing, Infrared, Radar/Millimeter Wave, Acoustic, Seismic, Photometric/Visible, Ultraviolet, Magnetic, Simulation, Background Characterization, and Operational Effectiveness

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Grant Gerhart
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	Unlimited	313	586-574-8634

NOTICES

Disclaimer

The findings of this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

The citation of commercial products or trade names in this report does not constitute an official endorsement of approval of such products.

Distribution Restriction

Distribution unlimited. For additional copies, refer requests to Signature Research, Inc., P.O. Box 346, Calumet, MI 49913, (906) 337-3360.

ACKNOWLEDGEMENTS

The Thirteenth Annual Ground Target Modeling and Validation Conference was, once again, a success this year because of the effort, dedication, and support of a number of people. The sponsorship of TARDEC was greatly appreciated. Host, Signature Research, Inc., is thanked for their assistance with local arrangements and essential support activities. The authors and presenters are acknowledged for the high technical quality of their presentations and papers, a vital contribution to making this conference a success.

<u>2002 Conference Executive Committee</u>

Dr. Grant Gerhart
U.S. Army TARDEC
Warren, MI

Mr. William Reynolds
Signature Research, Inc.
Calumet, MI

Mr. Dave Thomas
U.S. Army TARDEC
Warren, MI

FOREWORD

The annual conference on Ground Target Modeling and Validation is sponsored by the U.S. Army Tank-automotive Command Research, Development & Engineering Center (TARDEC). It is administered through Signature Research, Inc. (SGR). The conference is held on the campus of Michigan Technological University in Houghton, Michigan. The Thirteenth Annual Conference on Ground Target Modeling and Validation was held 05-08 August 2002.

The conference is held at the UNCLASSIFIED level and is open to all interested persons. In general, the attendance is made up of technical-level individuals representing industry, defense contractors, and government employees. The focus of the conference is on modeling and simulation of ground target signatures and backgrounds, and the validation of such models. Multi-mode technologies include the infrared or near-infrared, radar/millimeter wave, acoustic, seismic, photometric/visible, ultraviolet, and magnetic sensing.

The 2002 proceedings will be provided to all attendees as part of the symposium registration. Contact Signature Research, Inc. at (906) 337-3360, or **gtmv@signatureresearchinc.com** for additional copies of these proceedings.

Table of Contents

Thirteenth Annual
Ground Target Modeling and Validation Conference Proceedings

Session VI: Signature and Sensor Modeling
Session Chair: Mr. David Thomas, US Army TACOM

Session VII: Man-in-the-Loop Performance
Session Chair: Mr. Gary Witus, Turing Associates, Inc.

Design and development of a snapshot imaging spectropolarimeter

Derek S. Sabatke, Ann M. Locke, Eustace L. Dereniak,
John P. Garcia, Christopher P. Tebow

University of Arizona, Tucson, AZ 85721

David Sass

US Army TACOM, Warren, MI 48397

ABSTRACT

A snapshot imaging spectropolarimeter is under development. The instrument operates through the synthesis of the techniques of computed tomography imaging spectrometry (CTIS) and channeled spectropolarimetry. CTIS is a technique of imaging spectrometry in which an image of a scene is captured through a computer-generated holographic disperser. Spatially resolved spectra are reconstructed from the diffracted orders in post-processing. Channeled spectropolarimetry is a method of encoding the four components of a spectrally varying Stokes vector into a single irradiance spectrum. The combination of the two techniques provides the basis of an imaging spectropolarimeter capable of acquiring a complete data set in a single focal plane array integration time with no moving parts.

Our description of the instrument includes a review of the basic principles of both techniques. The diffraction orders generated by the CTIS disperser amount to projections through the scene's object cube, which provide an analogy to the medical imaging technique of computed tomography. Channeled spectropolarimetry is analogous to sideband modulation (as may be used in radio communication). Thick (high order) retarders and a polarizer are employed to introduce modulation. This viewpoint provides a starting point for the formulation of system requirements and reconstruction techniques. The requisite spectral resolution of the embedded spectrometer and selection of carrier frequencies by design of the retarders are addressed. We also discuss the design of a prototype snapshot imager for the visible spectrum and describe a non-imaging spectropolarimeter for the development of calibration and data reduction techniques.

1 INTRODUCTION

Imaging spectrometry and polarimetry are promising techniques for the detection and identification of manmade objects in natural backgrounds. Spectrometry enables detailed comparison of target and background spectra. Polarimetry allows the observer to capitalize on the fact that light emitted from and reflected by smooth surfaces tends to acquire a polarized component.[1] Polarization information then has the potential to highlight manmade objects despite spectral camouflage. We report on development of an imaging instrument which integrates both spectrometry and polarimetry functions. The instrument will be capable of characterizing the state of polarization of radiation from each pixel of a target scene by measuring all four components of the Stokes vector as a function of wavelength.

The data acquired can be interpreted as an image of a four-dimensional volume, since a measure of radiance is obtained for four independent variables or indices: two spatial variables (x, y), wavenumber (σ), and the Stokes vector index (j). It should be noted however that the Stokes vector index has only four possible values (the integers from 0 to 3), whereas the x, y, and σ dimensions will each be segmented into a greater number of intervals. We refer to this four-dimensional volume as the spectropolarimetric hypercube. For example, our goal for a proof-of-principle

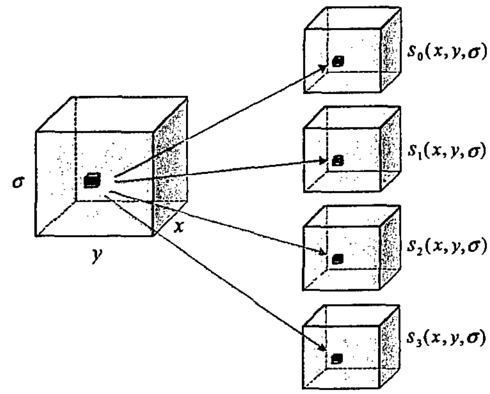


Figure 1: An illustration of the four-dimensional (x, y, σ, j) nature of the data acquired by an imaging spectropolarimeter.

system for the visible spectrum is to obtain on the order of 50×50 pixel spatial resolution with 16 wavelength bands in the spectral image for each Stokes component. Figure 1 illustrates the concept of the four-dimensional hypercube.

Conventional spectrometers and polarimeters are generally unable to image all of the dimensions of the hypercube at once. They are inherently sensitive to a one-, two-, or three-dimensional subset of the volume, and must scan out the remaining dimensions in some manner. For example, a camera with a narrow-band filter could be used to obtain a single x,y slice through the hypercube. In this slice x and y vary while σ is fixed at the wavenumber passed by the filter and j is held at zero. The entire four dimensions could be swept out by swapping in filters with different transmission wavelengths and sets of polarizers and retarders, using two independent filter wheels in front of the camera. Similar examples can be made for other systems, such as slit spectrometers and whisk broom scanners (see Figure 2).

Two drawbacks of systems which require scanning are immediately apparent. First, moving parts (such as rotating filter wheels and dithering mirrors) are generally employed, and these are undesirable from the standpoint of reliability. Second, a relatively long time is required for the capture of a complete data set, since multiple exposures are made sequentially in time. Changes in the target scene during scanning manifest themselves as artifacts in the results. Scanning systems thus have limited ability to acquire data on rapidly changing scenes, such as moving targets or targets viewed from moving platforms. A brute force method for avoiding scanning is to use several separate systems viewing the same scene in parallel, possibly with beam splitters to direct light to each. Such systems tend to be expensive, however, since multiple focal plane arrays (FPAs) and a considerable investment in optics are generally involved. Furthermore, registration of the images from the separate systems is problematic.

The system under development promises to circumvent these difficulties via snapshot capability. It will capture data with coverage of the entire hypercube in a single integration time with a single FPA.

1.1 Channeled spectropolarimetry

The instrument will operate through the fusion of two techniques: channeled spectropolarimetry[2–4] and computed tomography imaging spectrometry.[5] Accordingly, it is referred to hereafter as a computed tomography imaging channeled spectropolarimeter (CTICS). Figure 3 illustrates the principle of operation of a channeled spectropolarimeter. The radiation under analysis is passed through two thick (high order) retarders and a polarizer, and the spectrum of the exiting light is recorded by a spectrometer. The fast axis of the first retarder is aligned with the transmission axis of the polarizer, and the second retarder is oriented with its fast axis at 45° to the polarizer's axis.

The recorded spectrum is a linear superposition of the Stokes component spectra of the incident light, in which the coefficients are sinusoidal terms depending on the retardances of the retarders. Since each retardance is nominally proportional to wavenumber σ, the Stokes component spectra are modulated. With proper choice of modulation

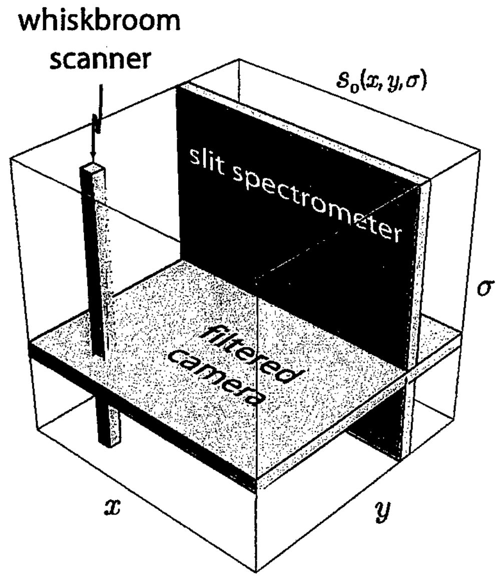


Figure 2: The lower-dimensional volumes which various spectrometer types are capable of imaging without scanning. Such spectrometers could be used with scanning to obtain spectral and spatial data on the s_0 polarization component, with additional measures (such as swapping of polarizers and retarders) necessary to obtain complete polarimetric data.

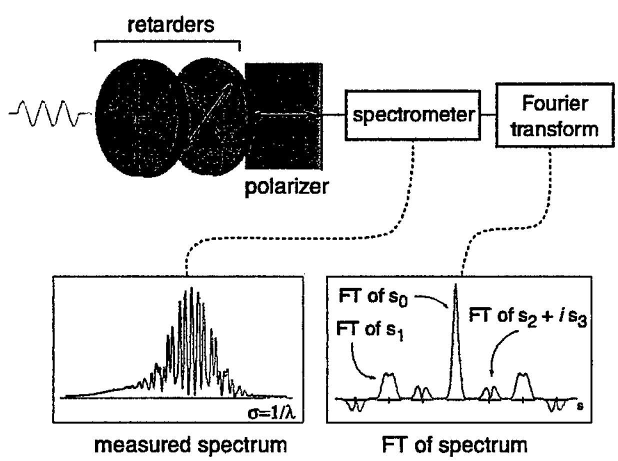


Figure 3: The channeled spectropolarimeter (after Oka and Kato[2]). The complex spectrum recorded at the output of the polarization optics is formed by a superposition of the Stokes component spectra modulating carriers. With proper choice of carrier frequencies, the Stokes components can be isolated in the Fourier domain.

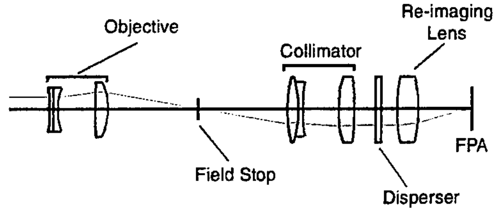


Figure 4: The CTIS Layout. The lens depictions are meant only to be representative of the role each lens plays in the system. Commercial objectives are usually used in the roles of objective, collimator, and re-imaging lens.

frequencies (i.e. proper choice of retarder thicknesses) the Stokes component spectra can be separated in the Fourier domain. This technique is analogous to sideband modulation in radio communications.

The term *channel*, as used in the name given to this technique, occurs as the convergence of two different meanings. In the sideband modulation analogy, individual spectra may be regarded as channels (in analogy to television channels). Dark bands or nulls in an optical spectrum are also termed channels, and a spectrum containing them is sometimes referred to as a channeled spectrum.[6] The carrier waves in the channeled spectropolarimeter result in channeled spectra, as illustrated in Figure 3.

1.2 Computed tomography imaging spectrometry

If used in combination with a snapshot imaging spectrometer, the channeled spectropolarimetry technique acquires snapshot imaging capability. The computed tomography imaging spectrometer (CTIS) is just such a spectrometer. CTIS obtains spatial and spectral data simultaneously by imaging through a computer-generated holographic disperser and carrying out a reconstruction using the mathematics of limited-angle tomography.

The basic layout of the CTIS architecture is shown in Figure 4. The objective lens forms an image of the scene under study in the field stop, which defines the instrument's field of view. The light from this image is then collimated, passed through a computer-generated holographic disperser (CGH) and imaged onto a camera's FPA. The image from each wavelength present in the scene is dispersed into a grid of diffraction orders on the FPA, with the separation between orders increasing with wavelength.

Imaging spectrometers gather data over a three-dimensional (x, y, and λ) volume, sometimes referred to as the object cube. The object cube is a subset of the four-dimensional hypercube discussed earlier. The effect of the dispersion of the CGH can be viewed as generating projections from various angles through the object cube onto the FPA, as illustrated in Figure 5. An estimate of the object cube is reconstructed from these projections by computed tomography. The reconstruction is quite computationally intensive and is carried out on a high-performance computer workstation. The spatial and spectral resolution achieved is constrained in part by the number of pixels in the focal plane array. As one would expect, increasing the number of pixels allows finer sampling of the spatial and spectral content of the image and therefore better spatial and spectral resolution. The complete CTICS system is obtained by combining the channeled spectropolarimeter and CTIS systems, as illustrated in Figure 6. Prototypes are under development for the visible and short wave infrared portions of the spectrum. The design examples given here apply to the system for the visible spectrum.

2 RETARDER DESIGN

Assuming each retarder is fashioned from a single crystalline material, cut with its optic axis parallel to the retarder's faces, then each retardance δ_k will follow

$$\delta_k = 2\pi d_k \Delta n_k \sigma. \tag{1}$$

The retarder's thickness is given by d_k and Δn_k is its birefringence (the difference in indices of refraction for its slow and fast axes). It is clear from Equation 1 that the retardance of each retarder will vary nominally in proportion to

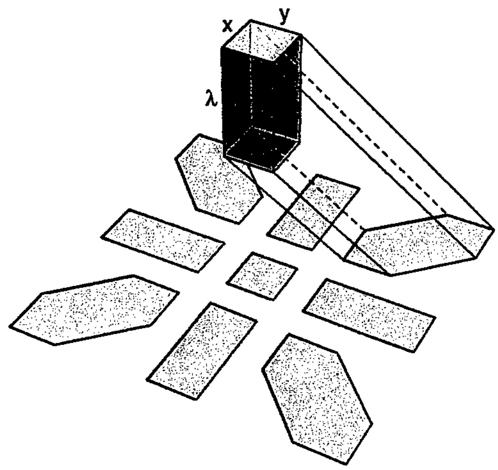


Figure 5: The dispersion of the CGH causes the separation of diffraction orders to increase with wavelength, resulting in several projections of the object cube onto the FPA.

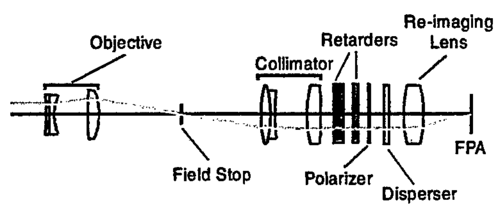


Figure 6: The CTICS Layout. The lens depictions are meant only to be representative of the role each lens plays in the system. Commercial objectives are usually used in the roles of objective, collimator, and re-imaging lens.

wavenumber σ. This dependence on wavenumber gives rise to the carrier waves in the sideband modulation analogy via factors of $\cos\delta_k$ and $\sin\delta_k$ which arise in a treatment of the system using the Mueller calculus. (See Sabatke *et al.*[7] for further details.)

The frequency of these sinusoids is set by the product of the retarder's thickness and birefringence, which may be recognized as the optical path difference (OPD) between the paths of rays polarized respectively along the slow and fast axes of the retarder. In many polarization applications, retarders are made as thin as possible in order to minimize the variation of retardance with wavelength. In the channeled spectropolarimeter, by contrast, the OPDs are deliberately designed with substantial values (which is why the retarders are referred to as *thick*). In real systems, the proportionality between retardance and wavenumber is imperfect because of dispersion in the material from which the retarder is fabricated. That is, the birefringence Δn_k in general depends on σ.

Some care is necessary in using the terminology of spectral analysis, since it can be applied in several ways to the channeled spectropolarimeter. Terms such as *frequency*, *band*, and *band-limited* apply to functions both of wavenumber σ and its conjugate variable under the Fourier transform, which we will denote s. In order to distinguish between the two meanings, we will often prepend σ or s to these terms.

Our system for the visible spectrum will operate over a wavelength range of 0.4–0.7 μm or a wavenumber range of 1.4–2.5 μm^{-1}. Our goal is to resolve 16 samples across the spectrum for each Stokes component spectrum. A 25% margin of extra s-bandwidth is designed into the system by using a value of 20 in place of 16. Then the OPD of the first retarder is calculated as

$$d_1 \Delta n_1 \approx \frac{20}{1.1\,\mu\text{m}^{-1}} \approx 18\,\mu\text{m}. \tag{2}$$

In order to obtain equally spaced channels (as illustrated in Figure 3) the retarders should have a 1:2 ratio in OPD. Thus the second retarder should have an OPD of about $36\,\mu$m. Dividing by the birefringence $\Delta n \approx 0.0096$ for quartz yields correspondng thicknesses of 1.9 mm and 3.8 mm respectively.

The tolerances on the thicknesses are fairly relaxed, since we are interested in setting the modulation frequency and not a specific number of waves of retardance. Tolerances of $\pm 0.5\%$ were specified to the vendor. The flatness of the retarder faces was specified at a quarter wave (at a test wavelength of $0.633\,\mu$m). It is possible for the thickness of the retarder to vary across its aperture as a result of a wedge angle between the faces. The result of such variation is that the carrier waves consist of a superposition of many sinusoids with slightly different frequencies. With increasing σ these sinusoids fall out of phase and begin to interfere destructively, resulting in a reduction of the carrier frequency amplitude.

For back-of-the-envelope purposes, we may assume that the frequencies which result from thickness variations in the retarder are uniformly distributed in a range Δh about a nominal carrier frequency h_{nom}. The carrier wave can then be represented as

$$\int \frac{1}{\Delta h}\,\text{rect}\left(\frac{h - h_{\text{nom}}}{\Delta h}\right) e^{i2\pi h\sigma} dh = \text{sinc}(\Delta h\,\sigma)\, e^{i2\pi h_{\text{nom}}\sigma}. \tag{3}$$

Thus the carrier takes the form of a sinusoid at the nominal frequency h_{nom} with its amplitude modified by a sinc envelope. As shown in Figure 7, the envelope tends to fall with increasing σ. Then if we dictate that the carrier amplitude must not drop below a specified amount (indicated by the dashed line in the figure), the value of the envelope at the upper wavenumber limit of the system's spectral band will be the key consideration. For example, to just meet a 95% amplitude requirement (that is, to have $\text{sinc}(\Delta h\, 2.5\,\mu\text{m}^{-1}) \approx 0.95$), a value of $\Delta h \approx 0.070\,\mu$m is required. This corresponds to a variation of $7.3\,\mu$m in thickness across the retarder's face. For a 1 inch diameter element, we arrive at a wedge tolerance of 60 seconds of arc. A margin of safety is allowed by specifying a tolerance of 15 seconds of arc. The retarders were fabricated with 1 inch diameter and a broad band anti-reflection coating on each face.

2.1 Non-imaging prototype

A non-imaging channeled spectropolarimeter has been assembled for demonstration of the method and development of calibration and data processing techniques. Retarders with a 3:1 ratio in OPD (rather than 1:2) are used. This is primarily a legacy of early treatments of channeled spectropolarimetry. The 1:2 ratio makes more efficient use of

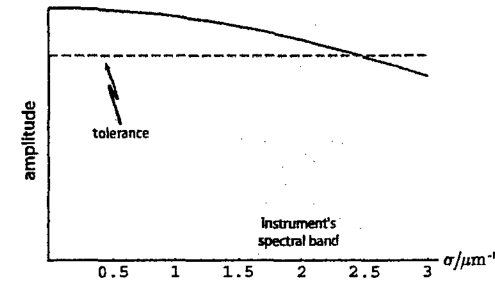


Figure 7: Variations in retarder thickness result in a roll-off of carrier amplitude across the spectrum.

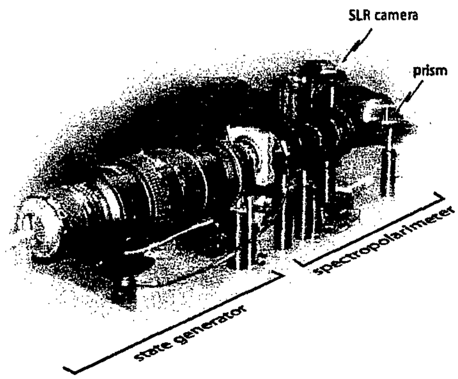


Figure 8: Experimental set-up for the visual demonstration of the channeled spectropolarimeter.

s-bandwidth in the spectrometer, and will likely be preferred for most purposes. However the principles of operation are the same for both configurations.

For quantitative measurements the beam exiting the analyzer is coupled into an optical fiber, and a fiber spectrometer is employed as the detector. A striking visual demonstration can be performed by replacing the fiber and spectrometer with a prism and SLR camera, allowing inspection of the channeled spectra with one's own eyes in the camera viewfinder. Figure 8 shows the system. A spectral polarization state generator, consisting of a white light beam sent through a polarizer and achromatic quarter wave plate, is used to control the input polarization state. Figure 9 shows spectra obtained with a similar setup (which differs in the use of a cylindrical lens to broaden the spectra in the direction perpendicular to that of the dispersion, and in the use of a digital camera to record the images).

3 CTIS DESIGN

The resolution required of the spectrometer is related to the desired resolution of the Stokes component spectra by a factor of seven (as can be seen by counting the s-domain channels in Figure 3). Since we are designing for 20 resolution bands in each channel, we require 140 resolution bands in the CTIS, giving

$$\delta\sigma_{\text{CTIS}} \approx \frac{1.1\,\mu\text{m}^{-1}}{140} \approx 7.9 \times 10^{-3}\,\mu\text{m}^{-1}. \tag{4}$$

The resolution of the CTIS is a complicated topic. The tomographic imaging method, calibration techniques and reconstruction techniques are all likely to impact its resolution. Practical estimates may be taken from the number of different wavelengths at which the system is calibrated and the number of FPA array pixels spanned by the CGH diffraction orders. Because CTIS is a grating spectrometer, any attempt to characterize its spectral resolution as a

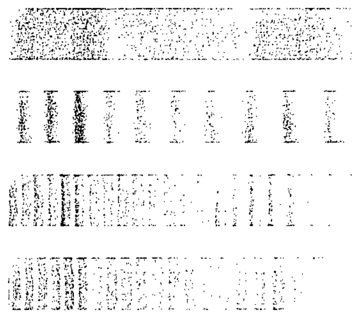


Figure 9: Reverse contrast images of channeled spectra for several different input polarization states, recorded with a system similar to that shown in Figure 8.

single value should probably be made in terms of wavelength rather than wavenumber. Wavelength resolution $\delta\lambda$ and wavenumber resolution $\delta\sigma$ are related (taking both to be positive) by

$$\delta\lambda = \frac{\delta\sigma}{\sigma^2}. \tag{5}$$

Assuming the CTIS wavelength resolution to be constant across the spectrum, its required value can be conservatively estimated by using the largest value acquired by the wavenumber in Equation 5. This yields $\delta\lambda \approx 1.3\,\text{nm}$, which corresponds to about 240 resolution bands in the CTIS. With this arrangement, the CTIS meets the wavenumber resolution requirement at the blue end of the spectrum and exceeds it (that is $\delta\sigma$ is smaller than necessary) toward the red.

In order to achieve 240 bands of resolution, the highest diffraction order in the CTIS (with a white input) should span at least as many pixels on the FPA. Table 1 gives the prescription for a CTIS system we have designed which meets this requirement. The prescription features a 2048×2048 pixel FPA (as shown in the left column of the table), collimating and reimaging optics with respective focal lengths of 300 mm and 60 mm, a 2 mm square field stop, and a period in the CGH of 16 μm (as shown in the right column). The system is intended to make use of a 5×5 array of diffraction orders (including the 0, ±1, and ±2 orders). The figure in Table 1 shows the footprint of the field stop after diffraction for each of the orders. The highest order is diffracted across more than 300 pixels, which meets our goal for spectral resolution. The direct (zero order) image of the field stop measures on the order of 50 pixels on a side. Hence we anticipate approximately 50×50 elements of spatial resolution and 16 elements of spectral resolution in each of the four Stokes components in the final CTICS system.

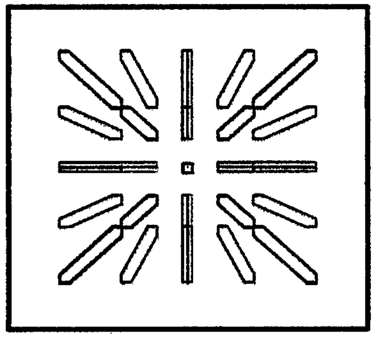
The CTIS prescription in Table 1 specifies the required period of the CGH. There is however considerable additional effort required to design a suitable CGH.[8] The CGH designer's task is to tailor the diffraction efficiencies of the orders so they increase modestly with order number in those orders used by the system (in this case the 0, ±1, and ±2 orders in both FPA dimensions), fall off sharply at higher order numbers to avoid sending light into unused orders, and remain well-behaved across a broad wavelength band. A phase-only hologram is used to preserve throughput.

The hologram is formed by etching a prescribed surface relief pattern into the face of a transparent substrate. The pattern is periodic, and is formed by replicating a smaller pattern (the unit cell) over the grating's face. The surface relief of the unit cell is described on a rectangular grid, with surface height taken to be constant within each grid element. Each element of the unit cell is termed a phase element (or *phasel*). For this instrument, the CGH unit cell was chosen to consist of an 8 × 8 array of phasels. Table 2 shows design data and the surface relief of the disperser designed for the spectropolarimeter. It was fabricated in PMMA using an electron beam etch.

Initial assembly of the CTICS has been carried out. Figure 10 shows a raw focal plane image. The object is a diagonal line in the field stop composed of white horizontal linearly polarized light.

Table 1: Prescription for the CTIS portion of the imaging channeled spectropolarimeter.

optical prescription	
field stop size / mm	CGH period / μm
2.00	16.00
collimator efl / mm	reimaging efl / mm
300	60
CGH aperture / mm	
19.0	

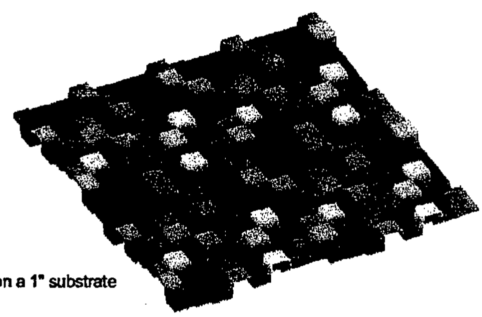


Focal Plane Array size	
x pixels	y pixels
2048	2048
x pixel size / μm	y pixel size / μm
7.4	7.4

wavelength range / μm	
low	high
0.4	0.7

Table 2: Design data for the holographic grating. The image is a computer's rendition of a 2×2 array of unit cells on the disperser's face.

material:	PMMA
index of refraction:	1.49608
maximum PV phase depth:	$1.65 * 2\pi$
design wavelength:	0.51 μm
performance band:	0.4-0.7 μm
Grating period:	16 μm
pixels per cell:	8x8
pixel size:	2 μm x 2 μm
distribution:	analog depth
full area:	19 mm active area on a 1" substrate



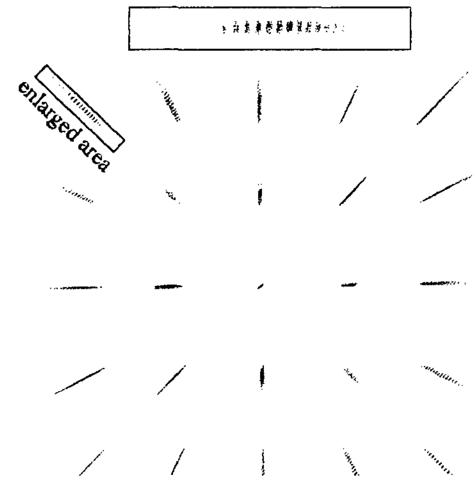


Figure 10: Reverse contrast raw image obtained at the focal plane of a CTICS system. The object is a diagonal line of white, horizontal linearly polarized light. A high diffraction order is shown magnified.

4 CONCLUSIONS

With the addition of a few polarization elements to the CTIS architecture, the spectral dependence of all four Stokes vector components is encoded in the measured spectrum. This synthesis of computed tomography imaging spectrometry and channeled spectropolarimetry holds great promise for a new imaging spectropolarimeter, CTICS. With a single focal plane array and snapshot capability, CTICS offers a novel solution to the limitations of conventional spectrometers and polarimeters. It is particularly well suited to the demands of rapidly changing scenes, as encountered in the identification and tracking of moving targets. We have discussed design considerations for CTICS instruments and are developing prototypes for the visible and short wave infrared. Ongoing work also includes development of calibration and data reconstruction techniques.

REFERENCES

[1] G. Forrsell, "Surface landmine and trip-wire detection using calibrated polarization measurements in the LWIR and SWIR," in *Subsurface and Surface Sensing Technologies and Applications III*, pp. 41–51, 2001.

[2] K. Oka and T. Kato, "Spectroscopic polarimetry with a channeled spectrum," *Opt. Lett.*, vol. 24, no. 21, pp. 1475–7, 1999.

[3] F. J. Iannarilli, S. H. Jones, H. E. Scott, and P. Kebabian, "Polarimetric-spectral intensity modulation (P-SIM): enabling simultaneous hyperspectral and polarimetric imaging," in *Infrared technology and applications XXV*, pp. 474–81, 1999.

[4] F. J. Iannarilli, J. A. Shaw, S. H. Jones, and H. E. Scott, "Snapshot LWIR hyperspectral polarimetric imager for ocean surface sensing," in *Polarization Analysis, Measurement, and Remote sensing III*, pp. 270–83, 2000.

[5] M. R. Descour, C. E. Volin, E. L. Dereniak, K. J. Thome, A. B. Schumacher, D. W. Wilson, and P. D. Maker, "Demonstration of a high-speed nonscanning imaging spectrometer," *Opt. Lett.*, vol. 22, no. 16, pp. 1271–3, 1997.

[6] R. W. Ditchburn, *Light*. London: Academic, 3 ed., 1976.

[7] D. Sabatke, A. Locke, E. L. Dereniak, M. Descour, J. Garcia, T. Hamilton, and R. W. McMillan, "Snapshot imaging spectropolarimeter," *Opt. Eng.*, vol. 41, no. 5, pp. 1048–1054, 2002.

[8] C. E. Volin, *Portable snapshot infrared imaging spectrometer*. PhD thesis, University of Arizona, Tucson, Arizona, 2000.

Synthetic Infrared Data For Target Identification Training and Testing

Bruce A. Weber, Joseph A. Penn
U. S. Army Research Laboratory
AMSRL-SE-SE
2800 Powder Mill Road
Adelphi, MD 20783

ABSTRACT

The performance of infrared (IR) target identification classifiers, trained on randomly selected subsets of target chips taken from larger databases of either synthetic or measured data, is shown to improve rapidly with increasing subset size. This increase continues until the new data no longer provides additional information at which point classifier performance levels off. It will also be shown that subsets of data selected with advanced knowledge can significantly outperform randomly selected sets, suggesting that classifier training-sets must be carefully selected if optimal performance is desired.

Performance will also be shown to be dependent on the quality of data used to train the classifier. Thus while increasing training set size generally improves classifier performance, the level of classifier performance improvement will be shown to depend on the similarity between the training data and testing data. In fact, if the training data to be added to a given set of training data is unlike the testing data, performance will often not improve but may possibly diminish. Having too much data can be as bad as having too little.

Our results again [1] demonstrate that an IR target-identification classifier, trained on synthetic images of targets and tested on measured images, can perform as well as a classifier trained on measured images alone. We also demonstrate that the combination of the measured and the synthetic image databases can be used to train a classifier whose performance exceeds that of classifiers trained on either database alone.

Results suggest that it may be possible to select data subsets from image databases that can optimize target classifiers performance for specific locations and operational scenarios.

Keywords: ATR, classifier, target identification, synthetic images, infrared

INTRODUCTION

Data available to train target-identification classifiers has long been known to be insufficient to produce robust target identification against new image datasets. This problem is due to the intimate relationship of data, the limited conditions under which data is collected, and the statistical nature of the classifiers in representing the data.

To remedy this, statistical classifiers need large amounts of dissimilar data for training. Obtaining measured data from field-tests is expensive so we have produced this data synthetically. To achieve this data synthesis we have created synthetic images that not only look like measured images, but moreover perform like measured images. To this end we use a comparison of measured, versus synthetic data trained classifier performance, as a quantitative measure of synthetic data validation.

Over time the performance of our synthetically trained classifiers has improved. Attention to detail in the comparison of measured and synthetic images was crucial. However, since our databases consisted of tens of thousands of images, a direct one-to-one image comparison was impractical. Instead we made our comparisons using the image-like states of a trained K-Means classifier. Such states, called codevectors or templates, are composite images that summarize many similar individual

image instances. This data compression makes practical the measured versus synthetic image comparison, and the subsequent adjustment and/or addition of new images and codevectors.

Last year we reported that synthetic-data trained classifiers could perform as well as measured-data trained classifiers [1]. However, as the subset of synthetic data was specially selected, we decided to investigate whether similar classifier performance could be achieved if randomly selected data was used to train the classifier.

Subsequently, we have expanded our synthetic database of four targets to slightly more than 90,432-files. From these images we have selected subset databases of varying size to train and test our classifiers. Two techniques were used to select subsets of synthetic image target-chips: (1) systematic unsupervised random selection to remove data bias and insure robustness to changing test conditions, and (2) biased selection using advanced knowledge optimized performance for limited conditions. The test database contained 5501-files of measured target-chips.

We next describe data creation, classifier training and testing, and data selection.

MEASURED IMAGE DATABASE EMULATION

For benchmark testing, we use the COMANCHE database of measured-world images. This database consists of approximately 30,000 image scenes containing different image instances of 10 different target types, 72 angular aspects spanning 360 degrees, three geographical locations including Yuma, Arizona, Hunter-Liggett, California, and Grayling, Michigan, both summer and winter seasons, and full diurnal time cycle. Extracted from these scenes are approximately 22,000 target chips that are divided into two databases: the SIG database of approximately 1,500 chips/target, including all 72-target aspects every 5°, for each of 10 targets in the clear, and the ROI database of approximately 1,300 chips/target, including only 8-target aspects every 45°, for each of 5 targets near clutter. All five of the ROI target types are included in the set of ten SIG target types. By any measure, the ROI target images are more difficult to recognize.

For comparison we have created synthetic data that emulates four of the ten targets in the SIG database. The four ground targets are: HMMWV, M60, T72, M113. Of these four targets, three are targets in the ROI database. We have modeled each target in identical conditions and locations, and have simulated realistic exercise routines to produce thermal signatures consistent with observed data. Similar exercise information is not available as part of the ground-truth for the measured data.

SYNTHETIC-IMAGE DATABASE GENERATION

Isothermal nodes for each target model were obtained using the PRISM [2] commercial code and IR-images were rendered using the Army Research Laboratory's (ARL) CREATION code [3]. Figure 1 shows a schematic of the algorithmic methodology [1] for generating, training, and testing of the synthetic target chip database. The dotted, blue Rhino-Muses [5] generation path provides a new methodology to be used in the coming year to replace the PRISM path used in this research for producing the isothermal target nodes.

For the purposes of this research, we quantify the validation of synthetic images by the performance to which a target identification classifier, trained on synthetic target chips, can achieve as compared to that achieved using measured target chips.

To create a synthetic database of ever increasing size we begin by selecting a small number of files from a large, parent database. This becomes database A. Next we add new files to create database B. Database B and each subsequent database of increasing size contain all the previous files selected for database A plus additional files. Each data selection is performed so as to uniformly represent the number of different targets but not necessarily uniformly represent their aspect or operating conditions. Files are randomly selected to form a database and then databases are enlarged, by the addition of more randomly selected files.

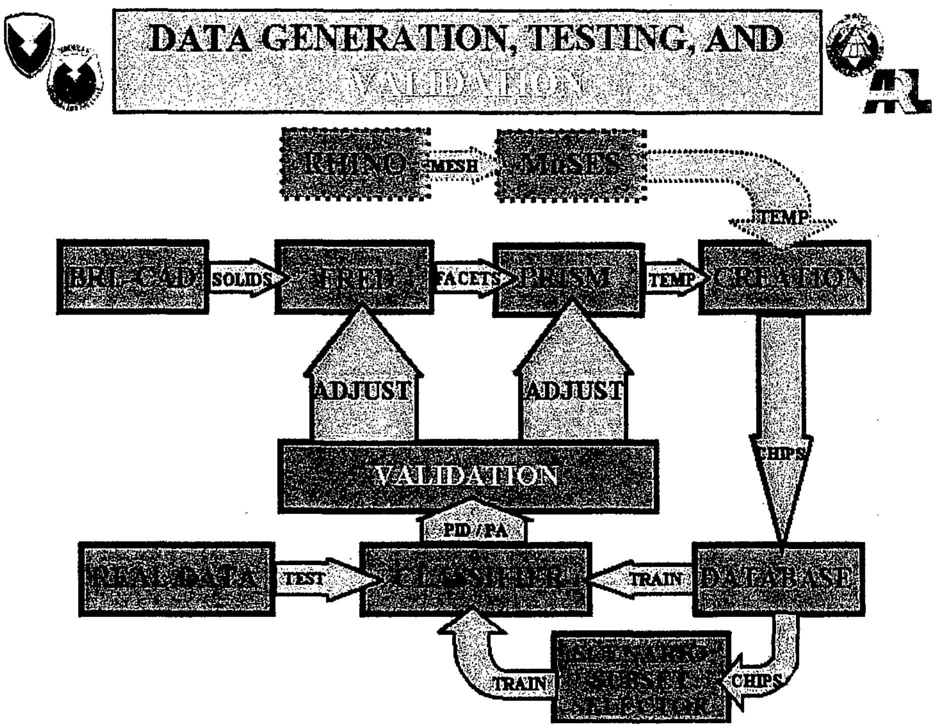


Figure 1. Iterative process for generating, training, and testing the synthetic image database.

As each database is formed a classifier is trained on the database and then the trained classifier is tested against the sequestered ROI database of measured-data.

No effort was made to adjust the ordering of the training data since the process of training averages like data into codevectors. Though training ordering does not affect a classifier's performance, database selection ordering is important with respect to database formation. (EXPLAIN THIS MORE CLEARLY!) Thus an intelligent or even lucky selection of a subset of data can outperform an unintelligently or randomly selected subset. Of course this is true only for subsets of large databases; when all of the data is used ordering does not matter.

CLASSIFIER TRAINING AND TESTING

The classifier used in this research was developed at ARL and is described as a minimized mean-squared-error (MSE) encoder [4]. All input target chips, both in the training and testing phase, are intensity scaled to zero mean and unity variance.

The classifier has two training modes: the K-Means mode, and the learning LVQ-mode. The LVQ mode is an additional mode that adjusts the results first obtained by the K-Means mode. In the K-Means mode a target-like region (identified in a previous step by a target detection algorithm) is extracted by a series of aspect dependent windows, enlarged to a fixed size, and wavelet decomposed into four (4) sub-bands. The K-means mode trains the classifier by collecting like-aspect sub-band decompositions and then creating codevectors by averaging the sum. The LVQ-mode adjusts the codevector centroids by moving the centroids to better match the training data. For the purposes of this paper we will be using the K-Means mode alone for classifier operation.

During the testing phase, an unknown target chip is extracted, sized, wavelet decomposed, and compared with each of the codevectors of each of the codebooks for each of the learned targets. The commonly used similarity measure Mean-Square-Error (MSE) is used for the comparison. The target-aspect sub-band class with the lowest MSE is declared the identity. This declaration is made independent of correct aspect identification.

DATA SELECTION

Scenarios and circumstance dictate data selection for target classification training. If a large enough database existed that contained images of targets in varying location, terrains, and seasons, then all of the data could be used to train a classifier to perform robustly in any condition. However databases are generally quite limited in the conditions they represent thus limiting the robustness of trained classifiers. Consequently scenario specific classifiers provide an alternate approach of achieving desired performance. Such classifiers are developed using advanced knowledge of a test site and/or test conditions.

Advanced knowledge might include data from recent reconnaissance, previous data collections from the test site or from sites similar in geophysical location and weather. But classifiers trained on advanced knowledge would have to be used selectively. Clearly one would not expect good performance from a classifier trained on data of targets in the Sinai in summer when the test site is Bosnia in winter.

To examine these possibilities we performed two tests: one in which synthetic data was selected randomly from a larger database to train a classifier and another in which the data was selected with advanced knowledge.

Random Data Selection

To demonstrate that synthetic data can be used to train classifiers and determine whether such training provides classifier performance comparable to classifiers trained on measured-data, we trained classifiers on either synthetic or measured-data, tested the trained classifiers on a sequestered set of measured-data, and compared the results.

Eight subset databases of increasing size were selected from the parent synthetic or measured databases. The selection was done so that any single database contained all of the data of any smaller database plus additional data. The data for each subset database, either synthetic or measured, was selected randomly. Equal percentages for each target in each subset was maintained relative to the larger, parent database, but the distribution of target aspects, seasons, times of day, and vehicle exercise states was not. The parent database for the measured-data contained 5,501-files and 90,432-files for the synthetic-data. Eight subset databases of 0.5, 1, 2.5, 5, 10, 25, 50, and 100% respectively were selected from either the synthetic and measured databases.

In addition, to testing classifiers trained on either synthetic or measured databases alone we also trained classifiers on combinations of synthetic and measured data. This was done so as to examine if overall performance could be improved by adding databases. If the data in the synthetic and measured datasets are similar then we might expect classifiers trained on any combination of data from the two sets to perform similarly. However, if the data in the datasets have some target chips that are not represented in the other set then combining data might improve the performance relative to either set alone. Of course any increase in performance would also indicate an increased similarity with the test dataset.

Randomly choosing data to form subset databases does not guarantee that the best performing groupings would be selected. In fact as increasing amounts of data dissimilar to the test database are added to a subset training-database some codevectors may be readjusted to be more unlike the test data degrading performance. This problem is addressed in the next section.

Advanced Knowledge Data Selection

Next we use a 3-step procedure to demonstrate that advanced knowledge can be used to increase classifier performance for targeted scenarios. First we divided a database of measured data into two equal sized subsets: a test dataset to be used solely for testing purposes, and a training dataset to be used for data selection. The training subset emulates our advanced knowledge in the form of previously obtained site data. This could be recent reconnaissance data from the test location, or data from another similar geophysical location. Next we trained a classifier with the training dataset, ran the entire synthetic database through the classifier and collected all files that were correctly identified. In the final step we used this selected database of synthetic target chips to train test our classifier. The dataset tested was the sequestered remaining half of the measured dataset not used to select the synthetic training dataset.

Two synthetic databases were selected in this way. One was chosen to emulate the ROI database and the other the SIG database. The ROI-Like set of synthetic data was selected using the ROI set of measured data, and the SIG-Like set selected using the SIG set of measured data.

For example the ROI-Like database was chosen by running the entire 90,432-file synthetic database through a classifier trained on the ROI training data. Selected ROI-Like images were required to be correct for both target identification and target aspect. They were also required to have a confidence value greater than 0.9. The confidence value is defined to be the difference in probability for correct identification between the first and second identification choice of the classifier. Using this process approximately 7,000 target chips were selected as the 3-target, ROI-Like database.

The SIG-Like database was chosen similarly. Using this process approximately 8,000 target chips were selected as the 4-target, SIG-Like database.

These two databases taken together with the 4-target, 5,156-image SIG database, and the 3-target, 2,200-image ROI testing database provides all of the datasets used to train and test our classifier.

RESULTS

Introduction

The two sections that follow describe our results: one addresses results obtained using randomly chosen training datasets and the other using datasets chosen using advanced knowledge.

For the testing described below, the trained classifiers are tested on the 3-target Measured-ROI-Testing set alone. Thus for performance in which Target 4 of the training set is not part of the testing set, no performance values are listed in column 4 since Target 4 can not be the correct identification. However this does not mean that Target 4 can not be declared to be the identity, only that it can not be the correct identity.

For the four-target test we are investigating, trained classifier targets 1 and 4, and targets 2 and 3 are of similar size, whereas targets 1 and 4 are smaller than targets 2 and 3. Target 1 is a HMMWV, target 2 is an M60, target 3 is a T72, and target 4 is an M113.

For all performance results reported here we used the K-Means mode of the LVQ-classifier. The K-Means parameters were chosen to optimize the number of codevectors generated for good target identification.

Results From Experiments Using Randomly Chosen Databases

In this section randomly selected training data is used for classifier training. Figure 2 shows four curves that summarize classifier performance using different sets of training data. The performance is specified as the probability of correct target identification (PID), in %/100, as a function of the percentage of the total number of training files added. For this graph a PID of 1.0 corresponds to all targets being identified correctly.

In Figure-2 the black curve (circles) shows the increase in performance of the SIG trained classifier as the percentage of the 5156-file dataset is increased. No PID saturation is observed since the PID never levels off. This indicates that each added increment of SIG files are sufficiently different from the SIG files used in smaller sets so as to add to the overall performance. The maximum PID achieved is 82%.

The blue curve (squares) shows the increase in performance for the synthetic data trained classifier as a percentage of the 90,432-file dataset is increased. PID saturation is observed as the PID begins leveling off for a training set of 9,000-files and is certainly level by about 45,000-files. Two possibilities may contribute to the PID saturation: (1) the new data being added is not sufficiently different to aid in the training, and (2) that the number of codevectors being created in the training process is not increasing. Figure-3 shows the increase in the number of codevectors as a function of the number of files per target is increased. Clearly there is no saturation in the number of codevectors being created as the database size increases. This suggests that the new data being added to the training databases is too similar to previously used files so as to not improve the performance of the classifier. The new data is redundant with other data in the overall 90,432 database. The maximum PID achieved is 70%. This suggests that randomly choosing data produces a smooth increase in performance, showing no evidence of subsets of data that outperform the total, and that using the entire synthetic database produces a level of performance 12% less than is produced by training the classifier on measured-data in the SIG database.

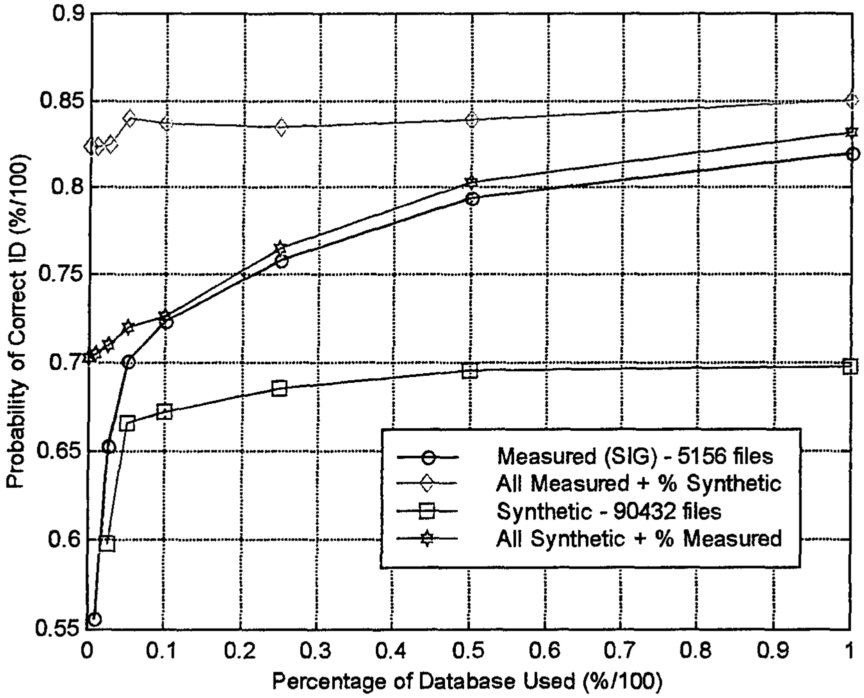


Figure 2 - Performance Comparison of K-Means Trained Classifiers Tested On Sequestered (ROI) Dataset of 5501-Files

The green curve (diamonds) shows the improvement in PID as increasing percentages of the 90,432-file synthetic database are added to the 5156-file SIG database. The performance is shown to improve to about a PID of 85%, a 3% increase over the performance observed using the SIG data alone. This difference is significant since the statistical uncertainty in recognizing files from the 5501-file ROI database is about 1.3%.

The violet curve (stars) shows the improvement in PID as increasing percentages of 5156-file SIG database is added to the 90.432-file synthetic database. The performance is shown to improve to about a PID of 83%, a 13% increase over the performance observed using the synthetic data alone and a 1% increase over the performance observed using the SIG data alone. This statistically insignificant difference shows that the SIG, and the synthetic + SIG trained classifiers perform essentially the same.

Clearly there is a difference between this end result here and that of the green curve (diamonds) for which one would expect the end points to be the same, that is, both trained on all of the SIG and all of the synthetic data. The difference here is the order in which the data used to train each subsequent classifier was selected. For the violet curve (hexagons) the classifier was first trained on all of the synthetic data before any SIG data was added. This biases the codevectors to be synthetic like. For the case of the green curve (diamonds) the classifier was first trained on the SIG data alone before any synthetic data was added. For this case the codevectors were first SIG like before they were adjusted by the addition of synthetic data. We conclude that the order in which data is added to the training process affects the final performance results achieved. We will see further evidence of this in the next section where datasets were selected using advanced knowledge.

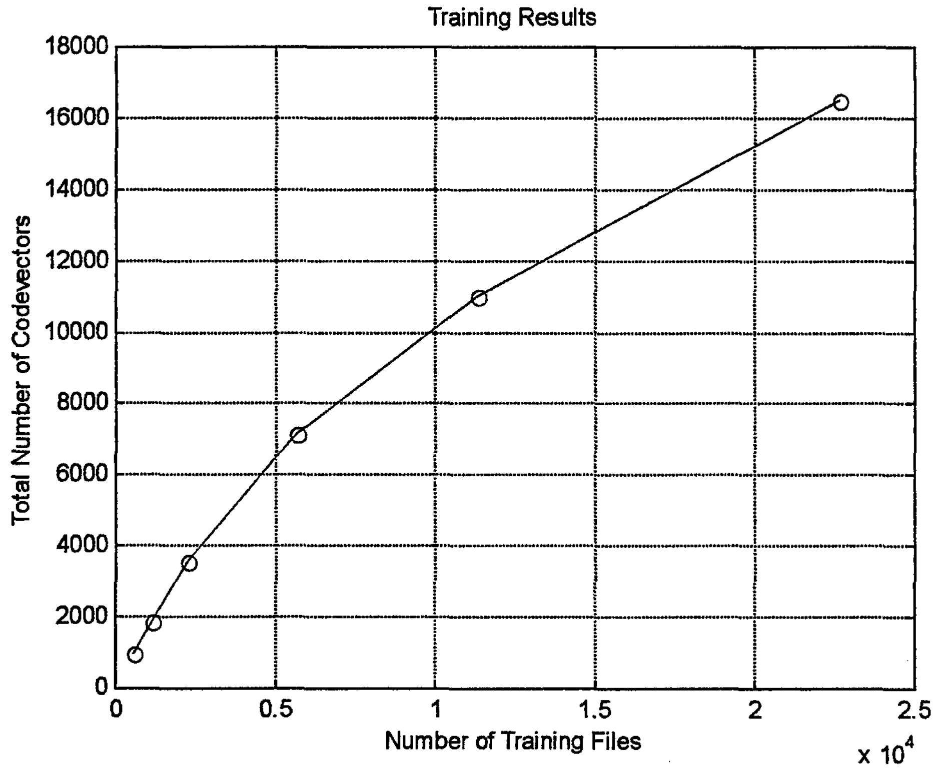


Figure 3 - This figure shows that the number of codevectors increases as the K-means classifier is trained on increasing amounts of synthetic data. The database included 22,608 target chips for each of four targets or 90,432 target chips in all.

Results From Experiments Using Advanced Knowledge

In this section advanced knowledge is used to choose training data. Tables 1 to 3 enumerate the identification probability (correct and incorrect) as a function of the target type. Such tables are termed a Confusion Matrix and the probabilities are shown in percent. The predicted target identities are listed down the left-most column and the actual identities are listed along the top-most row. The lighter, diagonal elements show the PID. A PID of 100% corresponds to all the targets of that type being identified correctly. Off diagonal elements represent the probability of misidentification.

Table 1 shows the confusion-matrix for the K-Means classifier trained on Measured-SIG data alone and tested on the sequestered Measured-ROI-Testing set. The overall PID is 82%. This is similar to the results reported by Chan et. al. [4]. Misidentifications of targets 2 and 3, both being tanks are often declared to be each other.

Probability of Correct Target Identification (%)				
Predicted/Actual	Target 1	Target 2	Target 3	Target 4
Target 1	85	4	6	-
Target 2	4	81	8	-
Target 3	2	12	81	-
Target 4	9	3	5	-

Table 1. A confusion matrix showing the probability for correct Measured-ROI target identification for a classifier trained on a database of Measured-SIG images. Overall probability for correct identification is 82%.

Table 2 shows the confusion matrix for the classifier trained on the union of the subsets of ROI-Like and SIG-Like synthetic data and tested on the sequestered Measured-ROI-Testing set. The overall PID is 81% indicating that the synthetic data trained classifier performs comparably to that of the Measured-SIG data trained classifier in Table 1. This is a different result

than that obtained when the synthetic-data training sets were selected randomly. In fact, comparing this result with that of the violet curve (stars) in Figure 2, we see that at no point does the PID rise above 70%, almost 11% below the PID obtained by choosing the training data using advanced knowledge. This suggests that using too much data that is unlike the testing data to train a classifier can bias the classifier performance in a non-optimal way.

Probability of Correct Target Identification (%)				
Predicted/Actual	Target 1	Target 2	Target 3	Target 4
Target 1	90	9	14	-
Target 2	5	76	9	-
Target 3	4	14	76	-
Target 4	1	1	1	-

Table 2. A confusion matrix showing the probability for correct Measured-ROI target identification for a classifier trained on a database of the union of the ROI-Like and SIG-Like synthetic subsets. Overall probability for correct identification is 81%.

Table 3 shows the confusion matrix for the classifier trained on the union of the Measured-SIG, SIG-Like, and ROI-Like synthetic datasets. The overall probability for correct identification is 85%, an increase of over 3% from the single database results for either classifier in Tables-1 and 2 above.. This demonstrates that classifier performance can be improved when synthetic and measured databases are joined. Again misidentifications mix Target 2 for Target 3.

Probability of Correct Target Identification (%)				
Predicted/Actual	Target 1	Target 2	Target 3	Target 4
Target 1	85	2	4	-
Target 2	4	85	10	-
Target 3	2	8	80	-
Target 4	9	5	6	-

Table 3. A confusion matrix showing the probability for correct Measured-ROI target identification for a classifier trained on the union of Measured-SIG, SIG-Like, and ROI-Like synthetic datasets. Overall probability for correct identification is 85%.

For comparison, Table 4 shows the confusion matrix for the classifier trained on measured data alone. The training data consists of the union of the Measured-SIG database and the Measured-ROI-Training subset. Again the test set is the Measured-ROI-Testing subset. The overall PID is 88%. This augmentation of the Measured-SIG data with the Measured-ROI-Training data increased the classifier performance by less than 9% over the results shown in Table 1. We will benchmark this level of classifier performance since it alone uses data taken from the same database from which the testing set was chosen.

Probability of Correct Target Identification (%)				
Predicted/Actual	Target 1	Target 2	Target 3	Target 4
Target 1	97	2	2	-
Target 2	2	85	15	-
Target 3	2	13	81	-
Target 4	0	0	1	-

Table 4. A confusion matrix showing the probability for correct Measured-ROI-Testing image identification for a classifier trained on a database of the union of Measured-SIG and Measured-ROI-Training images. Overall probability for correct identification is 88%.

Using the results of Table 4 as the benchmark the relative performance for the results listed for Tables 1, 2, and 3 are 93.5, 92.4, and 96.8 percent respectively.

CONCLUSIONS

Required target identification classifier performance specifications are dependent on application. Reconnaissance performance specifications can be considerable poorer than fire-control specifications. For specialized scenarios specialized classifier development will be required.

Statistical classifiers need a lot of data to train, but our results show that choosing data must be done carefully and wisely otherwise performance can suffer. Specifically we have shown that subsets of synthetic infrared-images can be chosen randomly to train target classifiers, and that adding synthetic databases to measured databases can improve the performance of classifiers trained on either database alone. We have shown that a classifier, trained with synthetic images selected using advanced knowledge, and tested on measured images, can perform as well as a classifier trained on measured images alone. Finally we have shown that using advanced knowledge to select training data, classifiers can be trained to significantly outperform classifiers that are trained on randomly selected training data.

Finally, we have achieved these results with relatively low-resolution images, derived from extremely low-resolution target models. We have taken care to simulate physically reasonable target states commensurate with measured data and we have validated our data by comparing synthetic to measured data performance in the training and testing of target classifiers. Yet our simulations do not include target/background interactions.

New updated infrared simulators, with near real-time temperature calculations and new visualization tools are now available [5], and soon high-resolution models will also be available. These tools will make database development easier and more reliable. And soon target and background thermal interactions will be modeled also.

Our methods for selecting data demonstrate that unless care is taken when choosing data, a range of performance is possible. Yet our methods do not provide a method on how to achieve optimal performance from available databases, but the way to proceed is clear. Choose data selectively using classifiers trained with advanced, scenario specific information from both measured and synthetic databases, train and evaluate the classifiers performance on available or reasonably matching measured data, and add new training data as it becomes available.

PRODUCTS

As a product of our work we have packaged our synthetic data. Datasets include the 90,432-chip, 4-target, COMANCHE-type ground target synthetic set, the 8,000-chip training SIG-Like dataset, and the 7,000-chip, testing ROI-Like dataset. In addition, along with each synthetic image dataset we are providing individual chip ground-truth as to geophysical location, time of day, month, weather, and vehicle exercise history.

These sets are unclassified and available (Distribution-C) to qualified U.S. Government Agencies and their contractors. Qualified users must agree not to distribute the data without first obtaining prior written approval from ARL.

ACKNOWLEDGEMENTS

We wish to extend a special thanks to Sandor Der of ARL for many helpful discussions.

REFERENCES

[1] Bruce A. Weber and Joseph A. Penn, "Comparable Performance For Classifier Trained On Real Or Synthetic IR¬Images," *Proc.SPIE AeroSense/Automatic Target Recognition XI*, Vol. 4379, April 2001

[2] Curran, A.R. Bobbi Wood, Eric Martilla, Paul Rogers, "Enhancements to the Vehicle Designer / PRISM Interface," *Fifth Annual Ground Target Modeling and Validation Conference*, Vol. 5, 401-407, August 1994

[3]Hung M. Nguyen, Joseph A. Penn, Teresa A. Kipp, Giap Huynh, and Janice Colby, "The CREATION Scene Generation Program And Its Application To Development And Testing Of Automatic Target Recognizers," *ATRWG Proceedings*, Vol. 1, 56-73, (July 1996).

[4] Lipchen A. Chan, Nasser M. Nasrabadi, and Vincent Mirelli, "Wavelet-Based Learning Vector Quantization for Automatic Target Recognition", *Proc. SPIE, Aerospace/Defense Sensing and Controls,* Vol. 2755, 82–93, April 1996.

[5] Johnson, Keith, Allen Curran, David Less, Derrick Levanen, Eric Marttila, Teresa Gonda, and Jack Jones, "MuSES: A New Heat and Signature Management Design Tool for Virtual Prototyping," *Proceedings of the Ninth Annual Ground Target Modeling & Validation Conference*, Vol. 9, 76-85, August 1998.

***Correspondence**: Email: bweber@arl.army.mil, telephone: 301-394-2501; and japenn@arl.army.mil, telephone: 301-394-1039; fax: 301-394-5234

Survivability of an Active Protection System during Combat

William Jackson, Jack Reed, Daniel Hicks
Survivability Optimization Modeling Team
Survivability Technology Area
US Army Tank-automotive and Armaments Command
Warren, Michigan 48397-5000

ABSTRACT

An Active Protection System (APS) consists of sensor(s), tracking radar(s), launcher(s), and countermeasure munitions. This technology is being postulated for the next generation of combat vehicles as well as a product improvement to existing vehicles. Active protection is believed to have a tremendous payoff by increasing the survivability of the ground combat vehicle without the burden of heavy armor. During combat, the components of the APS are subject to damage, which will degrade the performance of the APS. Using field data and engineering judgement, estimates of component damage from a single encounter are postulated for component packages of various sizes. This paper will answer the question: If an APS should last on average *k* rounds, then what size should the component package be?

INTRODUCTION

It is postulated that future ground combat vehicles will be much lighter than current ground combat vehicles. A consequence of this design trend is that armor, the traditional protection method, will only be a portion of the survivability solution for future ground combat vehicles. One of the techniques under both deployment and development, and the object of this paper, is that of an Active Protection System. It is postulated that an active protection system will reduce the need for armor. An AP system effects this reduction in armor by sensing an incoming threat, then tracking it, and at an appropriate time launching a counter-munition to intercept and destroy the incoming threat. Thus, the vehicles' armor need only deal with the residuals of the incoming threat, not the threat itself. There is, however, a fundamental difference between the protection afforded by traditional armor and that afforded by an AP system. Traditional armor is always effective (provided of course that one has enough of it), needs no action on the part of the user, and requires negligible amounts of maintenance. In contrast, an AP system may deplete its supply of countermunitions, or may suffer sufficient damage to one or more of its components that the system is rendered inoperable. The question that will be examined in this paper is the average length of operability for an AP system as a function of the survivability of its components.

STATEMENT OF THE PROBLEM

As indicated above, an AP system consists of various components: sensor(s), tracking radar(s), launcher(s), and countermunition. During battle these components are subject to damage, and damaged components will degrade the operability of the AP system so that it may be unable to meet its goal of vehicle protection. Component damage will alter the ability of the AP system to counter an incoming threat, but the lower the probabilities of component damage the longer the AP system can be expected to function effectively. The question this paper will address is: For a given level of component vulnerability, how long on average will the AP system continue to fully function.

BACKGROUND

The AP system considered in this paper is conceptually the same as the system considered in the paper *Functionality of Active Protection during Combat* by Caito, et al, [1]. Its four major subsystems, cueing sensors, tracking radar, launcher, and unguided countermunitions, are located outside of the vehicles armor envelope, and thus are subject to damage whenever the system encounters an ATGM.

The sequence of events when a functioning AP system engages a threat is as follows: The cueing sensor detects the launch of an ATGM and alerts the tracking radar. The tracking radar slews to the area of concern provided to it by the cueing sensor,

and begins tracking the threat, sending data to a decision-making-module (not considered a component of the AP system, since it has other functions and is under armor). The decision-making-module determines the path of the incoming ATGM, and thereby determines an intercept point at some appropriate standoff distance from the vehicle and a counter-munition launch time. After launch the counter-munition intercepts the ATGM, disperses its load of steel balls, thereby providing a cloud of steels balls through which the ATGM must pass. This passage will cause the ATGM to detonate, breaking the jet into fragments by the time it reaches the vehicle. These fragments will still hit the vehicle the AP system is trying to protect, but will not penetrate it. The fragments, however, may result in damage to the components of the AP system. A reduction in component damage will result in increased AP system functionality, and it is the investigation of this connection that is the object of the current paper.

From [1], the surface area of the various AP system components are:

Cueing sensors (2 sensors)	50 in^2
Tracking radar	144 in^2
Launcher with rockets	225 in^2

The fragment data from [1] assumed that the AP system's countermunition predetonated the ATGM at a standoff distance of fifty meters. The same fragment data showed that approximately 110 fragments from the dispersed jet hit the vehicle. However, since it is reasonable to assume, as was done in [1], that the AP system components are mounted around the periphery of the vehicle, the number of fragments impacting the component areas of the AP system were in the 11 to 33 fragments range. The potential component mounting area of the AP system considered here is that area outside a one-sigma area of the aim-point, and it is estimated to have a presented area of 80 square-feet.

It is assumed that if a single fragment impacts a component of the AP system, then that component becomes non-functional. That is, a worse case scenario is assumed. Thus, from the data in the preceding paragraph, it is possible to compute the probabilities that a fragment will hit a cueing sensor, a tracking radar, or a launcher/countermunitions combination as a function of the number the fragments impacting the vehicle's presented potential component mounting area, that is, the area outside one-sigma of the aim-point. This computation is summarized in Table 1 below.

Table 1: Probabilities of component damage during an ATGM engagement

Number of fragments	Prob hit, sensor	Prob hit, radar	Prob hit, launcher
5	0.01	0.06	0.09
10	0.02	0.12	0.18
15	0.03	0.17	0.26
20	0.04	0.22	0.33
25	0.05	0.27	0.39
30	0.06	0.31	0.45
35	0.07	0.36	0.50

The hit probabilities given in Table 1 are driven by the number of fragments impacting the potential component mounting area of the vehicle, and by the surface area of the exposed AP components. In [1], using the data given above, a model of an AP system's functionality was created using the theory of Markov chains (For background information on Markov chains, please see Isaacson and Madsen, [2], or Kemeny and Snell, [3]). Assuming thirty-fragment encounters, that model showed that after six encounters the AP system would be non-functional. More alarming, the model showed that the average duration of an AP system's functionality was only 1.55 encounters, and against less taxing ten-fragment encounters, the model showed that the average duration of functionality for the AP system was 3.41 encounters.

The goal of this paper is to construct a Markov model of an AP system similar to the one constructed in [1] and to use the model backwards. That is, given an average number of encounters for which the AP system is desired to be functional, what size must the component package (sensors, radar, etc.) be in order to achieve that goal? This answer of course will depend upon the number of fragments assumed to have hit the potential component mounting area of the vehicle.

DAMAGE MODEL

In the model of an AP system constructed here, it will be assumed that if the tracking radar and the launcher are both functional, then the AP system is functional. This is not an especially restrictive assumption, since if the sensors are damaged the radar can be left on to scan and track. In summary, the AP system is regarded as functional if both the radar and launcher are functional, and non-functional otherwise. Thus, the model constructed, a Markov chain, will have only two states: functional (F), and non-functional (N). Since it's impossible to go from state N to state F (that is, N is an absorbing state), the model is completely specified when the transition probability from state F to state N is determined. This transition probability, p_{FN}, will of course depend upon the presented surface area of the tracking radar and of the launcher, as well as the number of fragments assumed to be impacting the potential component mounting area (the area outside of one-sigma of the aim-point) of the vehicle.

Let r and l be the presented surface areas, respectively, of the tracking radar and the launcher (in square feet). Given that the potential component mounting area is 80 square-feet, the probability of a single fragment hitting the radar or the launcher is $p_1 = (r + l)/80$. Thus, the probability of a single fragment missing the critical components (radar and launcher) is $1 - p_1 = 1 - (r + l)/80$. If k fragments are assumed, then the probability of all k fragments missing the radar and launcher is $(1 - (r + l)/80)^k$, so the probability of at least one hit upon a critical component will be $1 - (1 - (r + l)/80)^k$. That is,

$$p_{FN} = 1 - (1 - (r + l)/80)^k.$$

This value completely determines the model, and thus, the transition matrix of the Markov chain of the model. The transition matrix is

Table 2: Transition matrix for APS model, r. l, k variable

	F	N
F	$(1 - (r + l)/80)^k$	$1 - (1 - (r + l)/80)^k$
N	0	1

ANALYSIS OF APS FUNCTIONALITY

In the notional AP system described above, r had a value of 144 in^2 or 1 ft^2 and l had a value of 225 in^2 or 1.56 ft^2, so that $r + l$ had a value of 2.56 ft^2. Furthermore, again from data discussed above, the worse case scenario for the number of fragments impacting the component area was 35 fragments. With these values the AP system model is

Table 3: Transition matrix for APS model, $r = 1.00$, $l = 1.56$, $k = 35$

	F	N
F	.320	.680
N	0	1

From this the average number of encounters for which the AP system will be functional can be computed by

$$\sum_{k=1}^{\infty} k(.68)(.32)^{k-1}$$

However, from Markov theory, see [2] or [3], this number, that is, the average number of encounters for which the AP system will be functional, can also be computed from $1/(1 - p_{FF}) = 1/.68 = 1.47$. This is not a stellar performance.

Since the AP system designer has no control over the number of fragments impacting the component area of the vehicle, better AP system performance must be obtained by reducing the size of $r + l$. Let $u = r + l$. Then, since k, the number of fragments, has been set to 35, the transition matrix of the AP system model is

Table 4: Transition matrix for APS model, variable $u = r + l$, $k = 35$

	F	N
F	$(1 - u/80)^{35}$	$1 - (1 - u/80)^{35}$
N	0	1

Now, the notional AP system described above in the BACKGROUND section carries four countermunitions. Thus, an appropriate first step in improving the performance of this notional AP system would be to reduce u enough so that the average number of encounters for which the system will be functional is at least 4. From above, it's known that the average number of encounters is given by $1/(1 - p_{FF})$. It follows that what is needed is the u satisfying the equation

$$\frac{1}{1-(1-u/80)^{35}} = 4$$

This gives $u = .655$ ft^2.

A reduction in the presented surface area of the tracking radar and the launcher from the current 2.560 ft^2 to 0.655 ft^2 may be a reduction so severe that it will be impossible to achieve despite component miniaturization. However, with miniaturization and hardening it may be that not all of the 35 fragments under consideration will have sufficient energy to damage the AP components. Hence, below are presented the values of u, $u(k)$, corresponding to a variable number of fragments, k, that will provide the AP system an average of four functional encounters.

Table 5: AP system functionality for an average of four encounters

k, number of fragments	$u(k)$, component presented surface area
5	4.473
10	2.269
15	1.520
20	1.142
25	0.915
30	0.763
35	0.655

From this table it's seen that the current notional AP system (2.56 ft^2 of presented surface area) would be on average fully functional for four encounters provided the number of fragments striking the component area were somewhere between 5 and 10 fragments. Furthermore, from the table it's seen that with miniaturization and hardening, it would not be unreasonable to believe that an AP system could be designed that would have an average functionality of four encounters.

Averages, however, can be misleading, especially in the model of an AP system constructed here. The reason is that the model assumes in the computation of an average number of encounters that any number of encounters is possible, despite the fact that there can be at most four encounters (for the notional system considered). It follows that a better question to ask is this: If the AP system is required with high probability to be fully functional for four encounters, then what size of presented component surface area, u, is required to accomplish that requirement.

Recall that $p_{FF} = (1 - u/80)^k$ and $p_{FN} = 1 - p_{FF} = 1 - (1 - u/80)^k$, where k is the number of fragments and u is the combined surface area of the radar and launcher, in the transition matrix, T, of the model

Table 6: General transition matrix, T, for the AP system model

	F	N
F	p_{FF}	p_{FN}
N	0	1

Now, if the vector $v = [f, 1-f]$ gives the probabilities of an AP system being functional, non-functional, respectively, before an encounter, then the vector

$$vT = [fp_{FF}, 1-fp_{FF}]$$

gives the probabilities of the AP system being functional, non-functional after the encounter, or, what is the same, before the next encounter. This computation uses the fact that $p_{FN} = 1 - p_{FF}$. Likewise, $vT^2 = [fp_{FF}^2, 1-fp_{FF}^2]$ provides the probabilities of the AP system being functional, non-functional after two encounters, and, in general, $vT^n = [fp_{FF}^n, 1-fp_{FF}^n]$ gives the probabilities of the system being functional, non-functional after n encounters. For more details, please see either [2] or [3].

The question posed above can now be answered. The original functionality vector, v, can be taken to be [1, 0]. That is, the AP system is fully functional. The probability that the AP system will be functional at the beginning of the fourth encounter is, as just noted, p_{FF}^3, since in this instance $f = 1$. What is desired is that this value should be high, say .95. That is the AP system will be at least 95% functional for four encounters. Thus, it is required that $p_{FF}^3 \geq .95$. Since $p_{FF} = (1 - u/80)^k$, where k is the number of fragments, the inequality to solve, u in terms of k, is

$$(1 - u/80)^{3k} \geq .95$$

The solutions of the above inequality for presented surface area of the radar and launcher, u, in terms of the number of fragments, k, are given in the following table. These are the values that will provide for the AP system remaining functional through four encounters with a probability of 95%.

Table 7: u required for four-encounter functionality with probability .95

u (square feet)	k
.273	5
.137	10
.091	15
.068	20
.055	25
.046	30
.039	35

These are miniscule numbers. The most favorable five-fragment scenario allows only 40 square-inches for the presented surface area of both the radar and launcher. It's difficult to see how this can be achieved even with minimization and hardening.

CONCLUSION

Given the current standoff distance of fifty meters for intercepting an incoming ATGM, it will be extremely stressing to design an AP system that will remain functional for four encounters with high probability. The standoff distance will need to be increased, but an appropriate standoff distance has not been determined. In order to estimate a workable standoff distance, additional field-test data needs to be accumulated. Meanwhile, the minimization and hardening of AP system components needs to be aggressively pursued.

REFERENCES

[1] Caito, Steven, Daniel Hicks, William Jackson, and Jack Reed, *Functionality of Active Protection During Combat*, Proceedings of the Eleventh Annual Ground Vehicle Survivability Symposium, held at the Naval Postgraduate School, Monterey, California, 2000.

[2] Isaacson, Dean L., and Richard W. Madsen, *Markov Chains Theory and Applications*, Robert E. Kreiger Publishing Company, 1985.

[3] Kemeny, John G., and J. Laurie Snell, *Finite Markov Chains*, D. Van Nostrand Company, 1960.

ModSAF-based development of DAS for light armoured vehicles

John Rapanotti* and Marc Palmarini†

*Defence Research and Development Canada – Valcartier, Val-Bélair (QC), G3J 1X5
†Onix Integration Inc., Ste-Foy (QC), G1V 4P1

ABSTRACT

Light Armoured Vehicles meet the requirement for rapid deployment by replacing passive armour with sensors, computers and countermeasures to detect and avoid threats. This emphasis on threat avoidance in an environment of increasing operational tempo results in major issues including: cognitive overload, training and skill decay and the significance of the man-machine interface. Modelling and simulation play an important role in these developments.

The integration of various technologies into a Defensive Aids Suite (DAS) can be designed and analyzed by combining field trials and laboratory data with modelling and simulation. ModSAF (Modular Semi-Automated Forces), is used to construct the virtual battlefield and, through scripted input files, a "fixed battle" approach is used to define and implement contributions from three different sources. These contributions include: models of technology and natural phenomena from scientists and engineers, tactics and doctrine from the military and detailed analyses from operations research. This approach ensures the modelling of processes known to be important regardless of the level of information available about the system. Survivability of DAS-equipped vehicles based on future and foreign technology can be investigated by ModSAF and assessed relative to a test vehicle. A system can be modelled phenomenologically until more information is available.

ModSAF is being developed for research and development, in addition to the original requirement of Simulation and Modelling for Acquisition, Rehearsal, Requirements and Training (SMARRT), and is becoming useful as a means for transferring technology to other users and researchers. This procedure eliminates the need to construct ad hoc models and databases. These concepts and approach will be discussed in the paper.

1. INTRODUCTION

Modern weapons have reduced the traditional effectiveness of passive armour on land vehicles. Portable missiles with warheads containing multiple shape charges can penetrate any thickness of armour. Sensor-fuzed munitions and top-attack missiles are designed to penetrate the more vulnerable turret. Artillery, instead of rocket motors, can be used to launch guided missiles that cannot be detected by missile approach warning systems searching for rocket plumes. The solution is therefore to avoid detection as long as possible by various means including camouflage and by reducing vehicle signatures to background levels. Survivability can be further increased by the early detection of threats followed by appropriate and timely countermeasures to either defeat the threat directly or to reduce the effectiveness of the guidance system.

A vehicle designed to survive these modern threats has to rely less on passive armour and more on sensors, computers and countermeasures. The long service life of the vehicle, typically 50 years, can also be a problem unless the vehicle is designed to accept upgrades. An approach to develop and maintain the survivability of the vehicle through a series of upgrades based on identified technological trends is discussed in this paper.

To better understand, evaluate and develop threat avoidance, realistic evaluations are carried out, in a context useful and relevant to the military, on a virtual battlefield. ModSAF (Modular Semi-Automated Forces) is used to construct the virtual battlefield and to model and simulate DAS[1–3] on light armoured vehicles. The DAS and LAV configurations being evaluated are described in more detail below.

Modelling physical systems in ModSAF is not new. Terrain features are represented in sufficient detail to study vehicle mobility, detection, defilade and other practical manoeuvres. Atmospheric phenomena are modelled to produce accurate effects of attenuation over distance, scattering by smoke and dust and insolation. Spectral effects in the atmosphere, such

Ground Target Modeling & Validation Thirteenth Annual Conference, August 2002

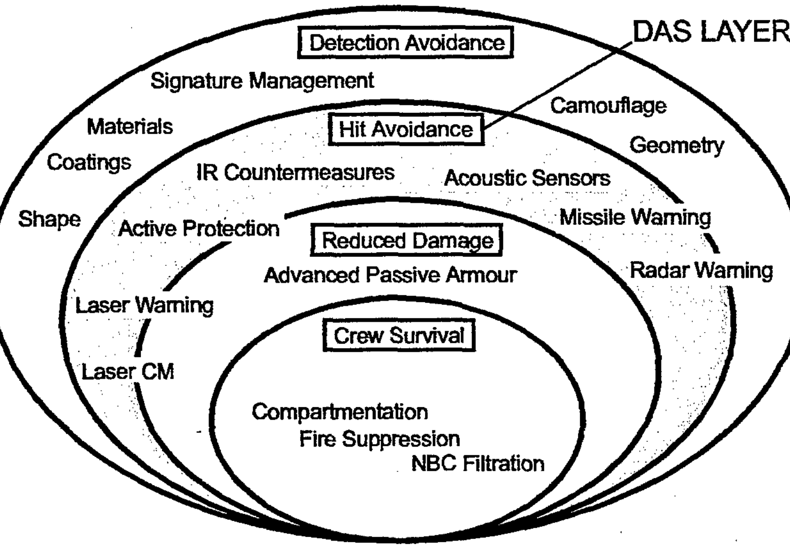


Figure 1. Layers of survivability. With the reduction of passive armour, greater emphasis is placed on detection avoidance and on hit avoidance, the DAS layer.

as propagation of artificial source in the solar-blind ultraviolet regime, natural effects such as solar glint and complicated, variable signatures from missiles are also modelled.

The combination of increasing computer power at low cost and the robustness of ModSAF can also be used to represent vehicles more realistically and in more realistic environments and evaluated more throughly than previously possible before final field evaluations. Any weapon system can be impoved by better materials and design. ModSAF can be used long before the system is fielded to develop new tactics and doctrine. Crew familiarization and training can be undertaken, initially, on stand-alone systems and progress through to vehicle simulators. In future vehicles, embedded simulators can be used to model the environment surrounding the vehicle, including terrain, atmosphere, threats and other vehicles. Some of the aspects of modelling a counterfire improvement based on a high-speed missile and a typical MBT countermeasure are discussed below.

2. SURVIVABILITY

Vehicle survivability can be represented usefully by a series of layers as shown in Figure 1. New vehicles designs place a greater emphasis on the first two layers, detection and hit avoidance, to survive an attack. In the first layer, survivability can be improved by reducing the size and silhouette of the vehicle and through signature management, which is the the reduction to background levels of the radar cross-section and signature in the visible, infrared, electronic, acoustic and magnetic domains.

The next layer, the DAS layer, relies on a system of sensors to collect data, which is then processed to determine the presence of any threats. This system is interfaced to countermeasures through processors, which will determine a prioritized list of responses. As the challenges of hit avoidance, including short timelines and numerous threats, are addressed, the solutions will lead to weapons of greater precision and faster tempo on the battlefield.

Among the many threats to land vehicles, a list of 89 missiles was compiled[3] in Table 1 according to the guidance and communication links used[4].

In this list, virtually all of the missiles have an operator in the loop leading to the possibilty of using a combination of dazzling and obscurants to disrupt the aiming sequence. An effective basic DAS could be based on laser detection, missile-detection tracking and countermeasures including dazzling, obscurants, counterfire and evasive manoeuvres. This "soft-kill" solution can be effective since the large number and variety of threat missiles can make identification, and therefore countermeasure selection, difficult. This difficulty can be overcome with a "hard-kill" solution, which will physically destroy the missile.

Table 1. Threat Missiles Classified by Guidance or Communications System

Number	Missile Type[4]
41	Semi-Automatic Command to Line of Sight (SACLOS)
16	Laser Beam Rider (LBR)
11	Manual Command to Line of Sight (MCLOS)
8	Fibre-optic guided missiles (FOGM)
7	Imaging Infrared
6	Laser and millimetric wave designation, including Semi-Active Homing
3	Laser based guidance or communications link
2	Automatic Command to Line of Sight (ACLOS)
1	Radio Frequency Homing
89/95	Total missiles/Total configurations

3. MODELLING AND SIMULATION

A Model-Test-Model cycle is difficult to establish for various reasons including, lack of information about foreign systems and incomplete models of the sensor and countermeasure environment. As shown in Figure 2, a continuous cycle can be established using field trials and experimental data to develop models and simulations. Ideally models should be based on physical principles but when this is impractical, systems can still be analyzed phenomenologically. Both approaches can be implemented in ModSAF. ModSAF (Modular Semi-Automated Forces) was developed for training and doctrine developement and provides a capability to define and control entities on a simulated battlefield. It is a model of the dynamic behavior of simulated units, their component vehicles and weapons systems with sufficient realism for training and combat development. ModSAF simulates an extensive list of entities including fixed and rotary wing aircraft, ground vehicles, dismounted infantry, and additional special models such as howitzers, mortars, minefields, and environmental effects. The behaviour of the simulated entities can be scripted so they can move, fire, sense, communicate and react without operator intervention. The entities can interact with each other as well as manned simulators, over a network supported by Distributed Interactive Simulation. Operating over a network is also useful in maintaining a necessary level of security.

These basic feature in ModSAF are sufficient to define the participation of three group of workers and implement their requirements free from mutual interference. To gain general acceptance, ModSAF development must meet the requirements of the scientists and engineers who develop the technology, the operations research community and the military developing tactics and doctrine. MATLAB®, which is designed for quick-prototyping and code generation, can be used for ModSAF development. MATLAB®modelling can also be used to share information with contractors and other researchers As shown in Figure 2, an important application of ModSAF is the generation of a battlefield environment for Man-In-the-Loop simulators. The MIL simulators are critical in the development of a suitable Man-Machine-Interface for the DAS.

4. DAS DEVELOPMENT FOR LAV

Rapid deployment of the vehicle to a wide range of possible missions and low cost upgrading plays a significant role in the design of the DAS. Some desirable DAS characteristics include a

fitted for, but not fitted with, approach providing a quick response at low cost implies designing the vehicles for equipment upgrades according to the mission requirements without needing to purchase for the entire fleet,

modularity including minimizing the interference among subsystems, which can complicate an upgrade and incremental upgrades of best of breed technology of a federation of modules instead of an integration of fused sensors,

mission configurability relying, for example, on the Galix grenade system that offer a wide range of capability from CS gas and stun grenades for peacekeeping to obscurants and fragmentation grenades for higher intensity warfare and a

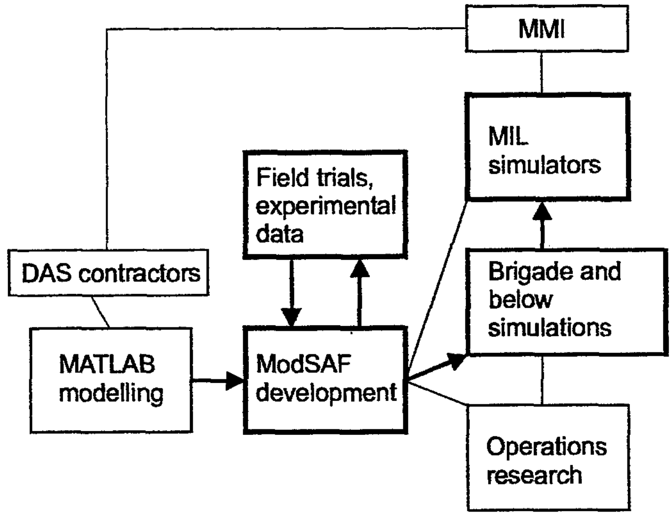


Figure 2. The four aspects of ModSAF development are shown. MATLAB ®is used as a quick-prototyping tool generating, transferrable models and code usable by ModSAF. There a tight loop between field evaluations and ModSAF development used to design DAS prototypes and plan future trials. Larger battles are carried out in simulation labs where new tactics and doctrine are developed. ModSAF is also used to provide the battlefield around Man-In-the-Loop simulators. From the simulators, the man-machine interface and vehicle operating systems are developed.

plug and play capability facilitating fast upgrading and replacement.

This level of readiness also facilitates rapid acquisition of up-to-date technology and further facilitates rapid deployment.

The DAS should be a federated, modular and mission configurable system, interfaced to the vehicle bus for access to other systems such as the Fire Control System. To keep the cost as low as possible the DAS based on more mature technology first and because of the rapidly evolving nature of technology modified through 5-year upgrades. DAS evolution is represented in Figure 3, could be carried out as described in the sections below.

During the 50year service life of the vehicle, 5year upgrade cycles will ensure peak performance at a reasonable cost. The basic vehicle configurations desrbied below do not preclude inclusion of other important systems such as sniper or bioaerosol detection and countermeasure. The Defensive Aids Suite is a group of sensors interfaced to countermeasures through data processors generating a list of prioritized responses. The DAS is also connected to the vehicle bus for access to the Fire Control System and the rest of the vehicle for counterfire and countermanoeuvres. The DAS also evolves through regular upgrades as shown in Figure 3. The 2010 and 2015 vehicles would be designed to operate in a network[5].

Present day LAV: The LAV has positive features including an electric turret with a slew rate limited to $\pm 45\,^{\circ}/s$ with a 25mm Chain Gun®with a range of about 2000m. The direct fire capability of the 25mm gun is improved with a laser rangefinder. The LAV has good mobility with a top speed of about 100km/h. The vehicle is however vulnerable to many threats and the ability to detect and counter threats is also limited. The vehicle defensive system includes grenade dispensers and a Laser Warning Receiver with limited one sector resolution. The smoke grenade is a NATO standard grenade with effective obscuration achieved in 2s and persisting for 30s. The metal flake composition produces a spectral coverage from visible to long-wave infrared.

Laser threat detection: a laser-warning receiver with an angular resolution of $\pm 22.5°$

Countermeasures: are carried out in sequence shown as required

- obscurants: VIRSS providing spectral coverage from visible to far infrared
- countermanoeuvres and

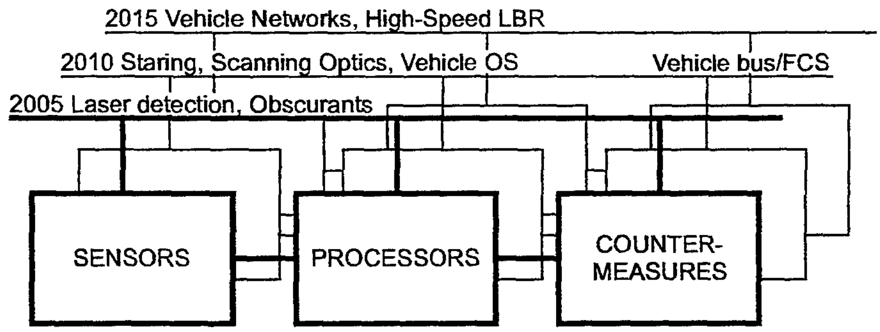


Figure 3. The rate at which computer and sensor technologies are developed justifies 5year upgrade increments. The more mature technology is implemented first beginning with laser-aided threat detection and visible/IR/MMW obscurants. Improved situational awareness, detection and identification is possible with staring and scanning optics. An operating system can be interfaced to the vehicle bus and fire control system providing automated response. By 2015, improved survivability can be achieved through vehicle networks and increased operational tempo with a high-speed missile.

- counterfire: based on the 25mm Chain Gun®and coaxial 7.62mm machine gun

2005 Light Armoured Vehicle: includes automatic, semi-automatic and manual launching of obscurants, manual countermanoeuvres and semi-automatic and manual counterfire. The Defensive Aids Suite will use the more precise HARLID™(High Angular Resolution Irradiance Detector) increasing the angular resolution [6] from one sector to $\pm 1°$. The P-MILDS will filter out most natural sources of UV and therefore can be used to detect artificial sources [7] including rocket propelled grenades, mortars and machine gun fire. Some missiles can also be detected and tracked where signatures have not been suppressed.

Laser threat detection: based on the HARLID™and the (BEam Rider Detector) to detect rangefinders, designators, beam riders and weapons over a hemisphere and an angular resolution of $\pm 1°$ up to a distance of 3km and 7m off-axis

UV threat detection: P-MILDS to detect artificial sources including rocket propelled grenades, mortars and machine gun fire over a hemisphere with an angular resolution of $\pm 2°$

Missile detection: based on Radar is used to improve threat detection and tracking

Countermeasures: are carried in sequence shown and as required.

- obscurants: VIRSS providing spectral coverage from visible to far infrared
- countermanoeuvres and
- counterfire: based on the 25mm Chain Gun®and coaxial 7.62mm machine gun

2010 Light Armoured Vehicle: The 2010 DAS is similar to the 2005 system including automatic, semi-automatic and manual response of counterfire, countermanoeuvres and obscurants. The optics used for detection and dazzling are depicted in Figure 4. Infrared Focal Plane Arrays provide a hemispheric coverage for increased situational awareness A system of organic UAVs is used to improve threat detection with a bird's eye view of the battlefield. Countermeasures are carried in sequence shown below, beginning with dazzling until full obscuration is achieved. Dazzling is intended to interfere with the operator by overloading his optical sight. Radar can be used with the infrared imaging system and smoke grenades to identify and defeat most threats.

High Availability (HA) principles are being used to develop reliable computer systems in critical applications and will probably influence the development of DAS.[5] The high level of reliability and transparency to the user will make the DAS much easier to accept. High Availability technologies available through Jini™include Alternate or Redundant Paths to Sensors, Dynamic Reconfiguration of the System comprising dynamic attachment and detachment and "hot pluggable" and "hot swappable" components.

- Real Time Operating System with Java™capability such as VxWorks®AE by Wind River Systems or LynxOS®by LynuxWorks™
- Computer architecture based on, VMEBus or CompactPCI™

The operating system is critical in the development of High Availability systems. Both VxWorks®AE, and LynxOS®have many of these features. VxWorks®AE is described as a RTOS with HA features including: Reliability, Availability, Serviceability, and Security (RASS).

Laser threat detection: based on the HARLID™and the BERD to detect rangefinders, designators, beam riders and weapons over a hemisphere and an angular resolution of $\pm 1°$ up to a distance of 3km and 7m off-axis

UV threat detection: P-MILDS to detect artificial sources including rocket propelled grenades, mortars and machine gun fire over a hemisphere with an angular resolution of $\pm 2°$

IR threat detection: Infrared Focal Plane Arrays for situational awareness including improved threat detection, 4096x4096 pixels per corner, providing hemispheric coverage, and a scanning NFOV system, typically $2.5° \times 2.5°$.

Visible threat detection: Laser illuminator and gated camera, typically $0.5° \times 0.5°$

Missile detection: based on Radar is used to improve threat detection and tracking

Remote sensing: with a system of organic Unmanned Aerial Vehicles

Countermeasures: are carried in sequence shown and as required.

- laser dazzling (requiring a directional platform),
- obscurants: VIRSS providing spectral coverage from visible to far infrared, chaff to extend coverage into the millimetre wave spectrum.
- countermanoeuvres and
- counterfire: based on the 25mm Chain Gun®and coaxial 7.62mm machine gun

2015 Light Armoured Vehicle: includes the capability of the previous vehicles. Radar is included for improved threat detection and tracking. The system of organic UAVs is replaced by a larger UAV with improved performance and providing extra information to the platoon of 4 vehicles. The vehicle network shares threat and countermeasure information with nearby platforms as shown in Figure 5.

Future vehicles can be configured into networks with an emphasis on the following networking technologies: [5]

- ISTAR network, to integrate all battlefield assets for command and control
- Jini™layer/Rio™facilitates a seamless and transparent integration of vehicles and platforms based on availablity, threat location and weapon capabilty.

The 25mm gun can be either replaced or augmented by a high-speed beam riding missile to increase the range to 5km with extra targeting information supplied by the UAV. Active armour is used to counter kinetic energy threats more effectively. The Radar system mentioned above is an essential component of the active armour system but may also be used directly to launch smoke grenades as required.

Laser threat detection: based on the HARLID™and the BERD to detect rangefinders, designators, beam riders and weapons over a hemisphere and an angular resolution of $\pm 1°$ up to a distance of 3km and 7m off-axis

UV threat detection: P-MILDS to detect artificial sources including rocket propelled grenades, mortars and machine gun fire over a hemisphere with an angular resolution of $\pm 2°$

IR threat detection: Infrared Focal Plane Arrays for situational awareness including improved threat detection, 4096x4096 pixels per corner, providing hemispheric coverage, and a scanning NFOV system, typically $2.5° \times 2.5°$.

Visible threat detection: Laser illuminator and gated camera, typically $0.5° \times 0.5°$

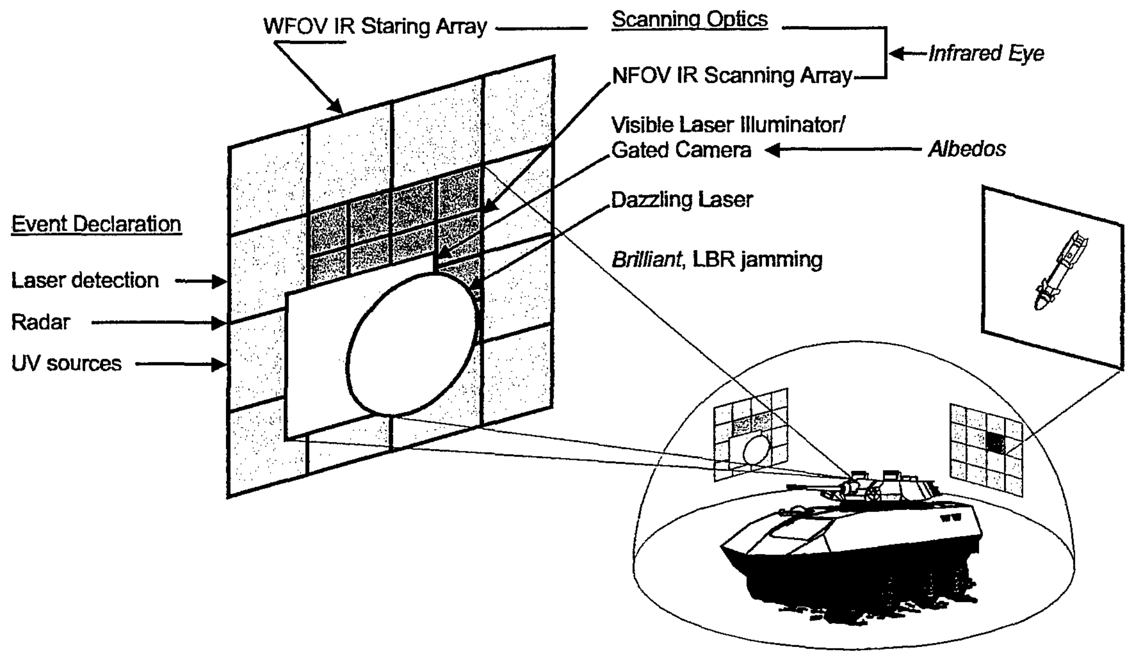


Figure 4. Most threats detected by the staring array *(far right)* will be subpixel in size. Detection can be improved by scanning with a higher resolution array *(center)*. A visible laser illuminator and a gated camera improves detection during nightime and low-light conditions. The visible and infrared imagery can be combined to provide a composite display for the crew. Dazzling disrpupts aiming and guidance while the main turret slew to position. Prototypes of these systems include: *Infrared Eye, ALBEDOS* and *Brilliant*, which is a laser beamrider countermeasure. Radar, UV imagers and laser detectors *(far left)* are used with these optical systems to identify potential threats. ModSAF will be used to analyze the timelines for detection, tracking and identification.

Missile detection: based on Radar is used to improve threat detection and tracking

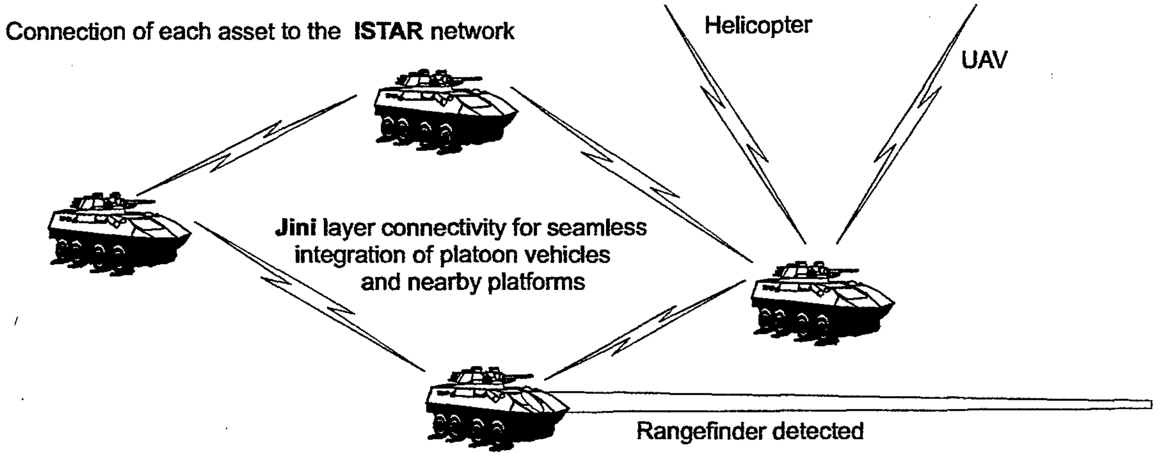
Remote sensing: with a UAV assigned to the platoon of four vehicles depicted in Figure 5.

Countermeasures: are carried in sequence shown and as required.

- laser dazzling (requiring a directional platform),
- obscurants: VIRSS providing spectral coverage from visible to far infrared, chaff to extend coverage into the millimetre wave spectrum.
- countermanoeuvres,
- counterfire based on
  - high-speed beam-riding missile,
  - 120mm low pressure gun,
  - 25mm Chain Gun®and coaxial 7.62mm machine gun
- active armour relying on the targeting information of the IRFPA and Radar.

5. MISSILE EVALUATION ENVIRONMENT

The direct fire capability of the LAV can be improved while maintaining a high operational tempo by developing a high-speed beam-riding missile. The missile is designed to accelerate a high density projectile, 0.5m long, to a velocity exceeding 2000m/s. The velocity is chosen to defeat soft-kill contermeasures by reducing the response time as much as possible and maintaining sufficient kinetic energy to penetrate MBT armour at the range limit of the main weapon. A typical engagement is shown in Figure 6. This missile is expected to be available for the 2015 vehicle described previously. A



SINGLE LAV	VEHICLE NETWORK
HARLID DETECTS RANGEFINDER AT 1.8km COMMANDER CHOOSES FROM APFSDS 150 ROUNDS HEI 60 ROUNDS APFSDS CHOSEN AND FIRED	HARLID DETECTS RANGEFINDER AT 1.8km *UAV CORRECTS RANGE TO 1.853km* COMMANDER CHOOSES FROM APFSDS *600* ROUNDS *ROCKETS 38* HEI *240* ROUNDS APFSDS CHOSEN AND FIRED

Figure 5. The UAV provides improved targeting information to the LAV Fire Control System. When a rangefinder pulse is detected. The commander in the single LAV, *left*, has to look down the direction indicated by the HARLID ™ to locate the threat. The HARLID™ in a vehicle network, *right*, detects the threat and shares this information with other platforms in the network. The UAV uses this approximate location to find threat and transmit a better estimate. The commander sees all the weapons available on the network including rockets from a helicopter hovering nearby. The APFSDS is selected and fired from any valid network gun. In the meantime, the commander protects the vehicle by dazzling, launching grenades and countermanoeuvring.

realistic evaluation would require that both the LAV and MBT be equipped with technology available for that time period. The MBT defence is based on the hardkill system, AWiSS-K, designed by DIEHL Munitionssysteme to stop kinetic energy penetrators.

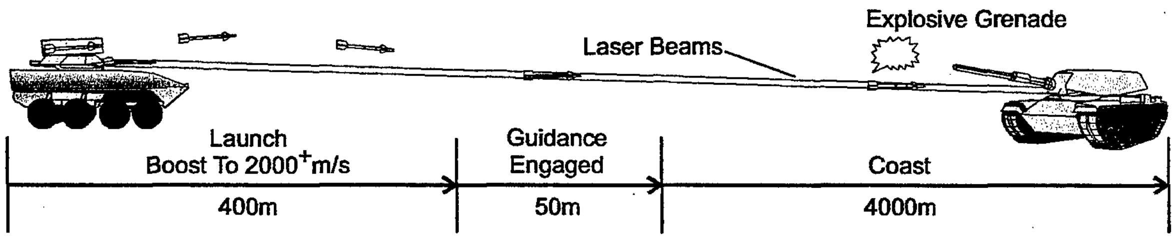


Figure 6. To keep the laser-based guidance path clear of propellant smoke, the missile is launched away from the beams until burn out occurs. The missile is then gathered in and guided to the target. The MBT *(right)* launches an explosive charge to deflect or destroy the missile.

5.1. Variables affecting missile lethality

The MBT and active protection will be modellied as follows:

Threat Detection and Tracking: The beat performance is expected from a combination of radar for velocity data and a high resolution staring array for more precise angular position.

Radar: limited to a range of 500m to avoid detection. The initial missile velocity and direction are measured to within $\pm$22m/s and $\pm$1millirad, respectively.

IR imagery: based on a system of infrared focal plane arrays with hemispheric coverage with effectively 4096x4096 pixels per corner. Detection algorithims will be used to alert the crew or automated system of MBT-like objects within range. It will be possible to detect and observe the orientation of the missile launchers and provide an early warning of a possible threat.

Countermeasure before missile launch: By detecting the turret-mounted launchers the MBT can react before the missile is launched.

Dazzling: Before effective obscuration occurs, which is about 2s for a NATO standard grenade, dazzling can be used to disrupt aiming.

Obscuration: Oscuration grenades based on a metal flake design can be used from visible through long wave infrared.

Radar and IR imagery provide an estimate of the missile position which is then used to aim the explosive charge. The AWiSS-K grenade is designed to explode at a fixed distance of 50m from the MBT launcher. It is possible to compensate for any systematic lag but there are still random variables that can affect the position and detonation of the explosive charge.

Countermeasure of launched missile: The DIEHL Munitionssysteme AWiSS-K is designed to deflect or destroy the missile. The explosive charge has a propagation velocity estimated at 2000m/s, but several variables can effect the positioning of the blast wave.

Grenade: variation in time to achieve maximum thrust, <10ms.

Explosive charge: variation in ignition lag time, <10ms.

6. CONCLUDING REMARKS

A procedure has been outlined to improve the development of DAS technology by combining prototype development and field trials with modelling and simulation based on operations research codes and off-the-shelf software tools. This new capability will provide a better estimate of vehicle performance on the battlefield and lower the cost of DAS development by complementing existing MIL facilities.

REFERENCES

1. J.L. Rapanotti, A. DeMontigny, M. Palmarini and A. Cantin, "Assessing the effectiveness of defensive aids suite technology", SPIE AeroSense Conference, April 2001.
2. J.L. Rapanotti, A. DeMontigny, M. Palmarini and A. Cantin, "Developing Vehicle Survivability on a Virtual Battlefield" , NATO Modelling and Simulation Group Symposium, November 2001.
3. J.L. Rapanotti, A. Cantin, and R.G. Dickinson, "Preliminary Study of Defensive Aids Suite Technology", DREV Technical Report (to be published).
4. R.G. Lee, T.K. Garland-Collins, D.E. Johnson, E. Archer, C. Sparkes, G.M. Moss and A.W. Mowat, "Guided Weapons", Bassey's (UK) Ltd., Third Edition 1998.
5. J.L. Rapanotti, M. Palmarini and M. Dumont, "New computer and communications environments for light armoured vehicles", SPIE AeroSense Conference, April 2002.
6. P. Webb, S. Soltesz, A. Cantin, J. Fortin and D. Pomerleau, "Improved miniaturized HARLID™ for laser warning systems having high angular resolution", SPIE AeroSense Conference, April 2001.
7. J.L. Rapanotti, "IR sensor design insight from missile plume prediction models", SPIE AeroSense Conference, April 2002.

Dazzle and obscuration strategies for light armoured vehicles

John Rapanotti* and Marc Palmarini†

*Defence Research and Development Canada – Valcartier, Val-Bélair (QC), G3J 1X5
†Onix Integration Inc., Ste-Foy (QC), G1V 4P1

ABSTRACT

Battlefield obscuration strategies, optimized for Main Battle Tanks in traditional high intensity conflicts, are inadequate when applied to Light Armoured Vehicles. LAVs are vulnerable to many threats and sufficiently different in design, capability and battlefield environment to benefit significantly from new strategies. Factors influencing this requirement include: i) the development of sensors with increasing accuracy and precision, ii) the need to minimize obscurant interference with vehicle sensors and other countermeasures, including active armour and explosive reactive armour, iii) the need to develop hemispherical obscurant coverage extending into the millimetre wave range, iv) grenades are needed to better match the increased tempo from greater vehicle speed, mobility and turret slew rate, v) the automatic configuration and selection of grenade burst patterns based on on-board processing and vehicle networks.

Spectral coverage in the visible to long-wave infrared regions is adequate, but trends in missile design are leading to the development of hybrid seekers including, laser designating, MMW seeking and imaging-infrared seeking capability accelerated by MEMS technology. With increased tempo, the time needed to achieve full obscuration becomes critical. Dazzling of a detected threat can be used to disrupt aiming and firing a second missile until full obscuration is achieved. Dazzling can also be used with the laser-illumination detection of optical systems. A generic threat response, based on dazzling and visible/IR/MMW grenades is preferred because of the large number of possible threats and the difficulty in developing practical identification strategies.

New dazzling and obscuration strategies, based on extensive knowledge acquired through field trials, will be analyzed and developed using ModSAF. These new strategies and the approach used to develop them will be discussed in the paper.

1. INTRODUCTION

Obscurants, dispersed by grenades, are an effective means of protecting the LAV against weapons using sensors for targeting and guidance.[1–5] Successful screening materials, such as metal flake and chaff, can reduce the effectiveness of anti-armour threats operating in the visible to MMW ranges. Brass flakes, typically 2-6μ in diameter, offers protection from visible to long-wave infrared, while chaff, consisting of aluminum coated fibres 10mm long and 25μ in diameter, is useful in extending coverage into the MMW range. Small particle dimensions are essential in developing a smoke screen that will remain suspended, or persisting, for the required 30s. Chaff dimensions, which can be relatively large to screen effectively nonetheless falls at an acceptable 0.3m/s or 9m in 30s.

Each grenade contains an explosive charge, which after a suitable time delay detonates to produce a cloud of uniform density. This cloud, approximated as an 8m sphere in this study, is actually an oblate spheroid aligned with the axis of the grenade and controlled by the launch angle and velocity of the grenade. Since the launcher is fixed to the turret, other variables affecting the launch include: vehicle pitch, roll and speed, turret position and turret slew rate. At low operating temperatures, the launch velocity is reduced resulting in a lower burst height. Once the initial momentum of the explosion has dissipated, atmospheric variables such as wind and turbulence distort and displace the sphere.

In peace-keeping roles, the grenade launcher will be an essential component launching a variety of grenades ranging from CS gas and illumination flares to fragmentation grenades. Unlike other platforms, land vehicles are relatively inexpensive and vulnerable to many threats.[6,7] These factors discourage the development of threat identification and favor a generic threat response like smoke screens. Since a grenade launcher will always be available, smoke screens will continue to play an important role in vehicle survivability.

Ground Target Modeling & Validation Thirteenth Annual Conference, August 2002

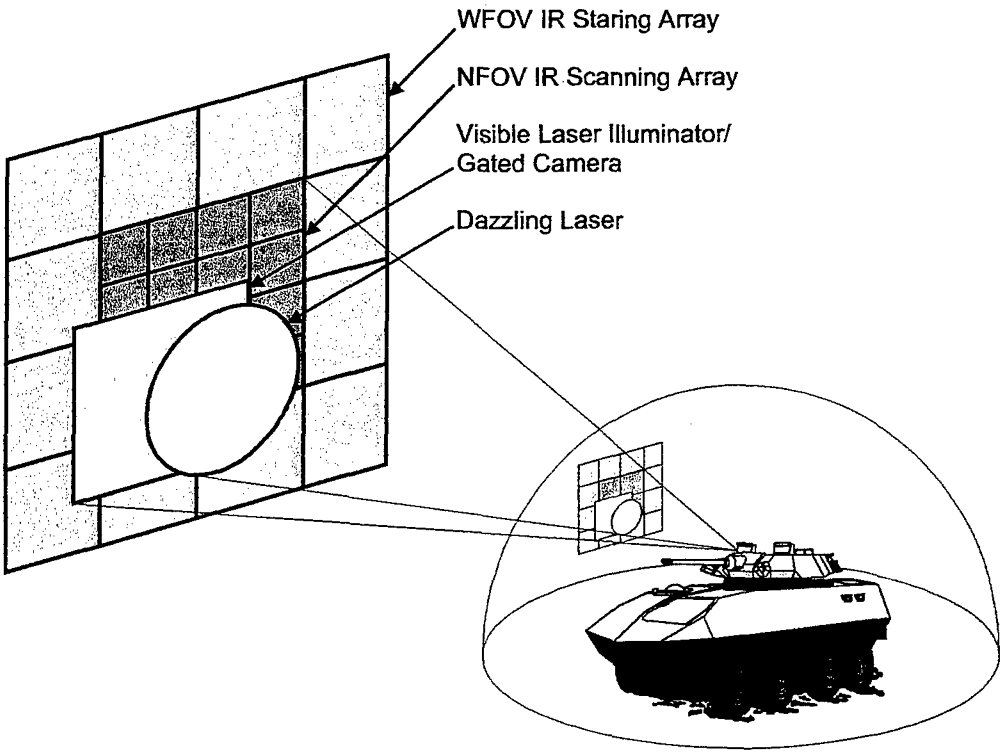


Figure 1. Vehicle staring and scanning optics. Most threats, such as missiles, are smaller than one pixel of the staring array. Detection can be improved by scanning with a higher resolution array. A visible laser illuminator and a gated camera can further improve detection especially during low-light conditions. The visible and infrared imagery can be combined to provide a composite display for the crew. Dazzling can be used, when appropriate, to disrpupt aiming and guidance while the main turret slew to position.

The interval between threat detection and full obscuration will be at least 1.5s. During this time, dazzling can be used to disrupt aiming or firing a second missile. The dazzling optics are a narrow field of view system housed in a mini-turret mounted on the main turret. Included in the mini-turret would be a laser illuminator and gated camera, ALBEDOS,[8] to actively detect various optical systems, by laser illumination, in the ALBEDOS field of view. The optics used for detection and dazzling are depicted in Figure 1.

The sections below will describe the factors influencing vehicle survivability and how dazzling and obscuration will be used to counter potential threats.

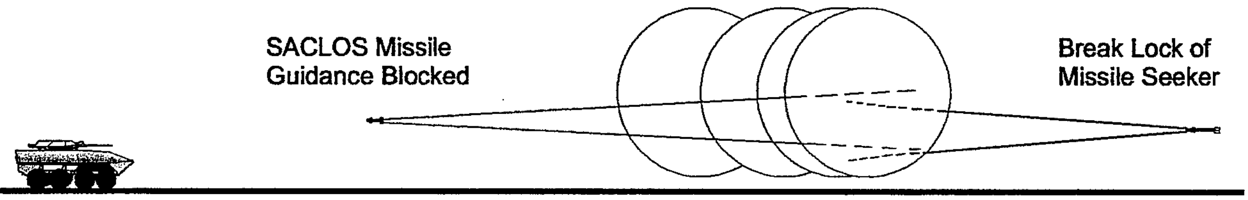
[illegible]

Obscuration screens are a practical means of defeating many threats by direct interference with targeting and guidance functions. Some factors influencing the use of obscurants with LAVs are discussed below.

[illegible] New generations of sensors are being developed providing greater levels of situation awareness. These performance improvements are being accelerated by MEMS technology to produce even smaller, hybrid systems with new properties based on combined characteristics. An example of a new detector is the laser detecting HARLID^{TM}. With an angular resolution of $\pm 1°$, it is a significant improvement over existing systems.[9] A current laser warning receiver with a typical resolution $22.5°$, can detect a threat but not provide the position with sufficient accuracy. The only reasonable response from the crew is to launch smoke grenades and back the vehicle away from the threat. Based on the HARLID^{TM}technology, a laser threat is detected in less than 1msec, but with a resolution $\pm 1°$ not accurately enough to position the main gun. Combined with an IR staring array, the stream of pixels corresponding to the laser source can be analyzed to determine the nature of the threat and fix the position. The information is then sent to the

Fire Control System and to other vehicles through a network.[10] With a staring array operating at 60Hz this process takes less than 20msec, considerably less than the typical 1.5sec it takes to set up sufficient obscuration.

[illegible] Obscuration over a wide spectrum can be used to defeat various missile systems including optically sighted, Semi-Active Command to Line Of Sight, and laser or MMW semi-active homing missiles. SACLOS missiles use a beacon facing the launcher to correct any deviations between the missile and the launcher crosshairs. Earlier designs were easily defeated by placing false beacons on the vehicle. These false beacons were much more powerful than the missile beacon and were used by the launcher to provide false trajectory data to the missile. Improvements in missile design, by encoding the beacon signal, resulted in a missile that could not be easily jammed. Both designs are susceptible to smoke screens, as shown in Figure 2, and can still be defeated by obscuring the flight path to the vehicle. The launcher no longer sees the target vehicle and the beacon signal is scattered and absorbed by the obscurant. Obsuration will also stop designated missiles since the laser or MMW beam cannot penetrate the smoke screen. New missile designs based on hybrid seekers: laser semi-active homing and both imaging IR and MMW imagery are being developed which will require careful manoeuvring forcing the missile to reacquire the target and correct trajectory over the distance between the vehicle and smoke screen.



[illegible] The LAV is protected by a screen formed by 4 grenades centered on a 36m radius. The smoke screen blocks the signal from the SACLOS missile guidance beacon. A missile seeker, initially locked on the vehicle, breaks lock and has only 32m to reacquire the target.

Obscurants designed to interfere with threat sensors will also interfere with vehicle sensors. A sufficient downrange distance is required to use active armour successfully. Careful selection and placement smoke screens is important in providing sufficient but not excessive downrange coverage. There is probably an optimum distance at which the smoke screen should be established, which can be determined through simulations with ModSAF.

[illegible] Light Armoured Vehicles will be deployed to peacekeeping environments where attacks can come from any direction. Sensors are being developed to provide the necessary hemi-spherical coverage but current grenade launchers, designed for Main Battle Tanks, need to be redesigned to provide a similar coverage. Improving sensor technology is also increasing the spectral range of weapons from visible and infrared to millimetre wave operation.

[illegible] Improved sensors and digital processing will automate many of the functions necessary in improving vehicle survivability. This automation with increased vehicle mobility and turret slew rate will shorten response timelines and increase operational tempo. The grenade launch velocity can be increased and the time delay shortened accordingly but the interval between threat detection and full obscuration will still exceed 1s. During this interval, dazzling is considered to be a reasonable countermeasure since most anti-armour threats rely on an operator to aim or guide the weapon.

[illegible] Obscuration will be set up according to the nature and location of the threat detected. This could be carried out automatically by Defensive Aids Suite processors based on local sensors or information transmitted over a network. The grenade burst patterns would depend upon threat detection and vehicle operation, described in detail below.

[illegible] The current MBT launcher has a 45° launch angle, which presents several problems. Any variation in the launch velocity, usually a function of the operating temperature, results in significant variations in the burst height. At very low temperatures, grenades often hit the ground before exploding. A second problem is the excessively long time delay, often in excess of 2.5s, required by the longer flight path. These problems can be avoided by

providing additional launch tubes at a shallower angle while retaining the 45 ° launch tubes for fragmentation grenades. Additionally, the shallower launch angle would be more appropriate for CS gas grenades.

[illegible]

A simplified governing equation including a given launcher angle, initial velocity and required launcher height can be expressed as:

$$h = h_o + V_o \sin(\alpha + \theta)t - 1/2gt^2 + V_s \sin(\alpha)t - V_o \sin(\alpha)t \tag{1}$$

where θ, is the launch angle (either 20°, 45° or 70°),
α, is the vehicle incline,
h_o, is the height of the launcher, set to 2.5m,
h is the height of the grenade, 4.3m at 20°, 18.0m at 45° and 26.7m at 70°, on flat ground
t is the time of flight, 1.5s,
V_o, is the initial grenade velocity, 20 and 25m/s,
V_s, is the vehicle velocity and
g, is acceleration due to gravity.

The burst pattern for the MBT, shown in Figure 3, can be improved by decreasing the grenade launch angle, increasing the launch velocity and shortening the time delay. Based on trials, the velocity is increased to 25m/s and the time delay is fixed at 1.5s. Solving for the burst height, for various launch angles and vehicle incline angles, results in a family of curves shown in Figure 4. For a wide range of vehicle inclines, the 20° angle gives the most acceptable distribution of burst heights. To maintain the requirement for fragmentation grenades, the 45° angle is retained for mid-level coverage. Further protection against top-attack weapons is provided with a single grenade at 70 °. A comparison between the MBT grenade system and the new LAV configuration is presented in Table 1. The total number of grenades has increased from 8 to 48 seems excessive but from previous studies[10] an automated system can be made more reliable if all the components are accessible by the computer. This implies installing all the grenades in the launcher instead of stored in the vehicle. The new burst pattern configuration for the LAVs is shown in Figure 5.

[illegible]

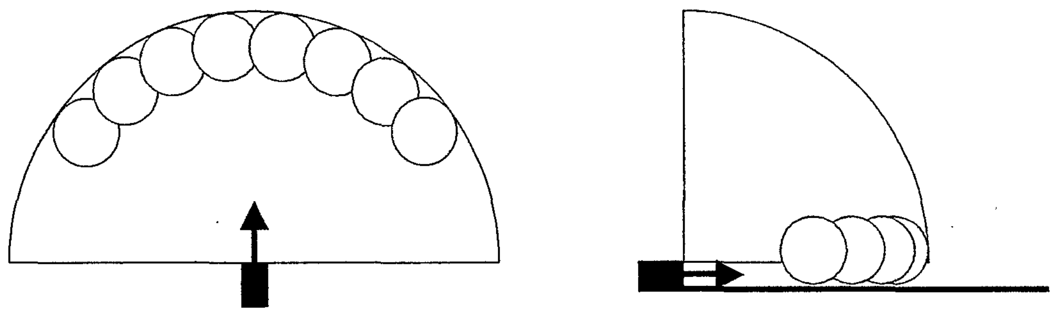
Based on the grenade configuration shown in Figure 5, various scenarios can be developed for further analysis. The objective is to automate the threat response as much as possible and reduce the crew work load.

[illegible] In the first scenario, a threat is detected while the vehicle is stopped or moving too slowly to avoid the threat. The recommended burst pattern is shown in Figure 6. The ground screen is formed with four grenades biased toward to rear so the driver can backup under cover. All three mid-level grenades including the 70 ° grenade and two aft mid-level grenades are used to counter a possible top attack. This allows the vehicle to back up and countermanoeevre for at least 30s. In a reasonably quiescent atmosphere, the 45° and 70° grenades should provide coverage well beyond the 30s required.

[illegible] For a moving vehicle, which is less vulnerable to sensor-fuzed submunitions, the burst pattern in Figure 7 is suggested. Both ground and mid-level grenades are used to form a series of screens, biased in the direction of vehicle travel. This procedure can be automated by launching the next set of grenades when the angle between the vehicle and the last grenade in the series approaches the angle of the threat detected. While this ensures that the vehicle remains hidden, it may still be possible to locate the vehicle by extrapolating grenade trajectories back to the launcher. If the driver, intentionally slows down or stops the vehicle, the variation of the scenario described above would be used provide protection while backing up.

[illegible]

The detection of threats by the staring array, the time to slew the scanning optics towards the threat and the time to slew the main threat are some of the stochastic variables that influence the usefulness of dazzling as a countermeasure. As suggested by Figure 8, if the time to slew the dazzling laser into place is excessive then the advantage over launching grenades may be negligible. Dazzling can be used preemptively with the scanning optics shown in Figure 1. Automatic processing can be used to quickly detect any anomalies against the background.



Dispersion - plan view

Dispersion - elevation view

Typical grenade-burst pattern for a Main Battle Tank. Each grenade explodes close to the ground forming 8m diameter spheres. A total of eight grenades are launched at 45 ° forming a smoke screen about 45m wide, 30m from the vehicle. The LAVs are expected to operate in very different threat environments requiring new strategies.

A new grenade launcher system, based on the requirements of Light Armoured Vehicles, is described. This approach is suitable for further analysis with wargaming simulations. ModSAF will be used to determine the best grenade configuration by constructing virtual battlefields and simulating vignettes based on accepted tactics and doctrine. Vehicle simulators will be used to develop man-machine interfaces and analyze vehicle and crew performance.

The grenade configurations described meet the LAV requirements of:

i) improved sensors,

ii) minimized obscurant interference,

iii) hemispherical screening from visible to millimetre wave range,

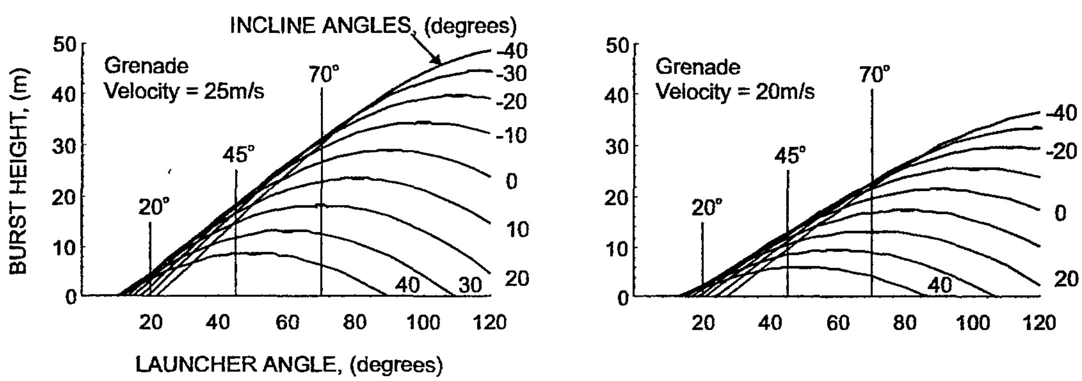
iv) increased operational tempo,

v) automatic configuration, selection and response.

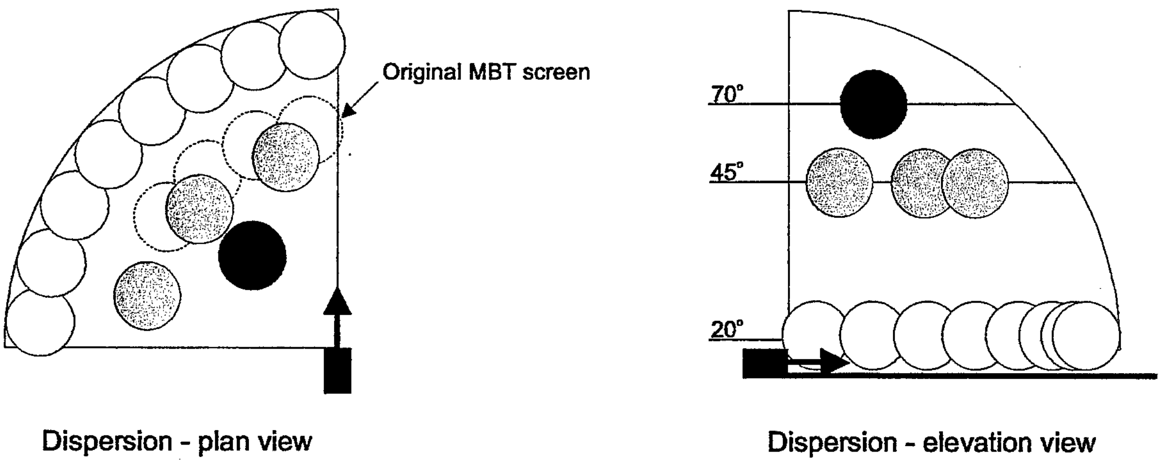
The original 45° launch tubes are retained for fragmentation grenades and new launch tubes at 20° are available for CS gas grenades.

[illegible] Main Battle Tank and Light Armoured Vehicle Grenade System Parameters

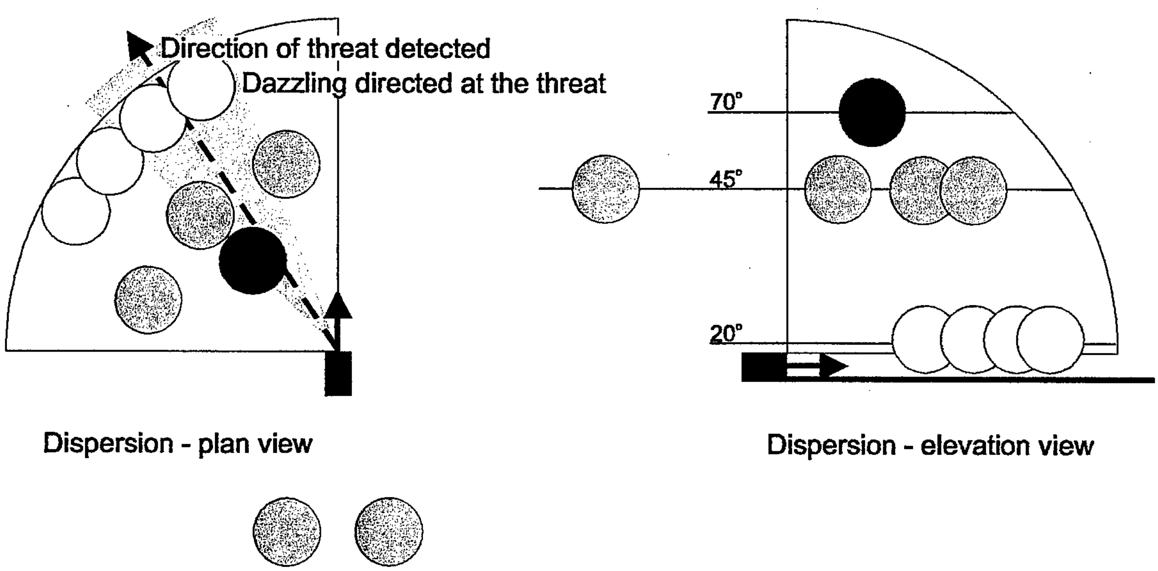
[illegible]	[illegible]	[illegible]
[illegible] [illegible]	Metal Flake Visible/IR	Metal Flake/Chaff Visible/IR/MMW
[illegible] [illegible] [illegible]	2.5s (approx.) 8m 30m	1.5s 8m 40m
[illegible] Ground level Mid level Top	[illegible] 8 launched at 45° — —	[illegible] 32 launched at 20° 12 at 45° 4 at 70°



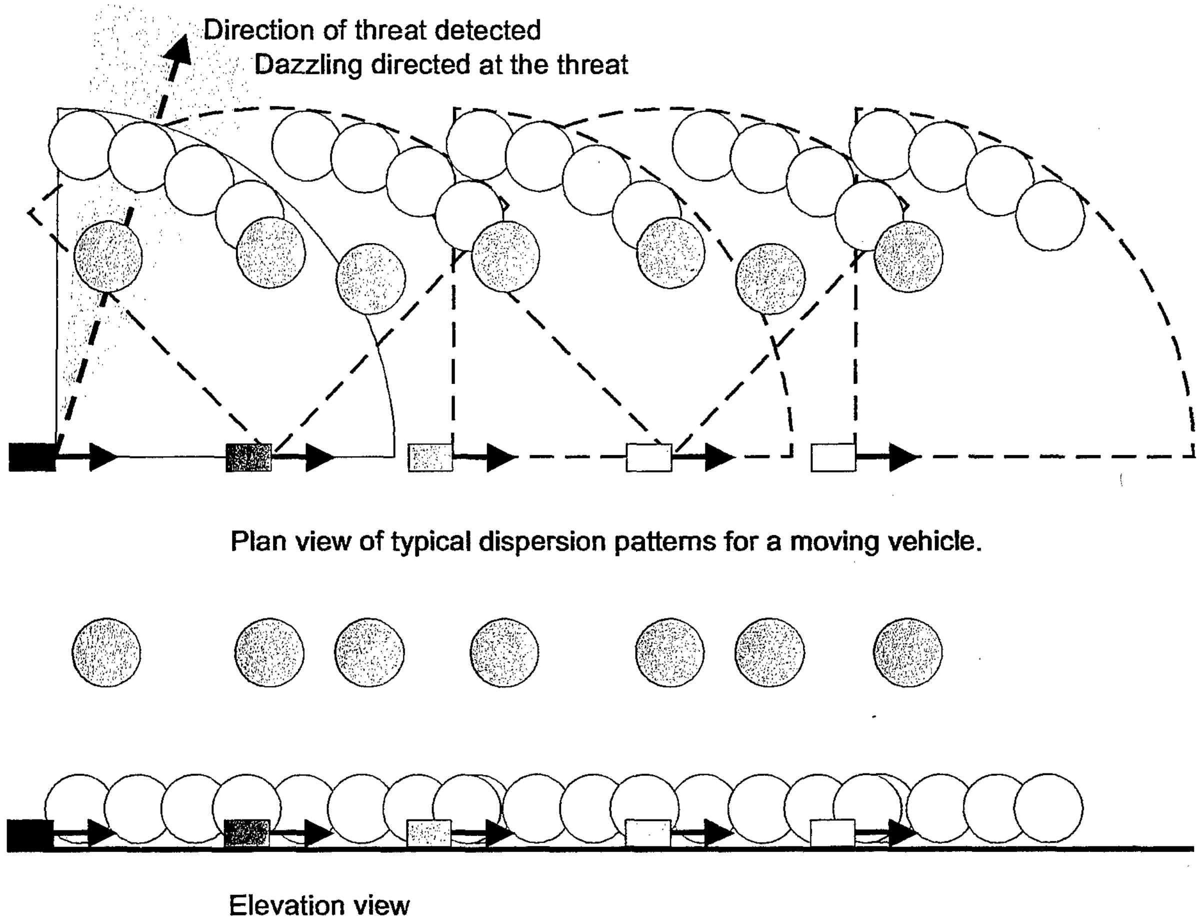
[illegible] Solution of the launcher equation for various launcher and vehicle angles. The effects of cold-environment operations are represented by launches at 20m/s. For incline angles from -40 ° to 40° most grenades explode before hitting the ground. The grenade at 70° would rarely be needed unless optimum coverage is required for a stationary vehicle. Other parameters include a delay time of 1.5s, a grenade initial velocity of 25m/s, a vehicle forward speed of 4m/s (14.4km/hr) and a launcher height of 2.5m.



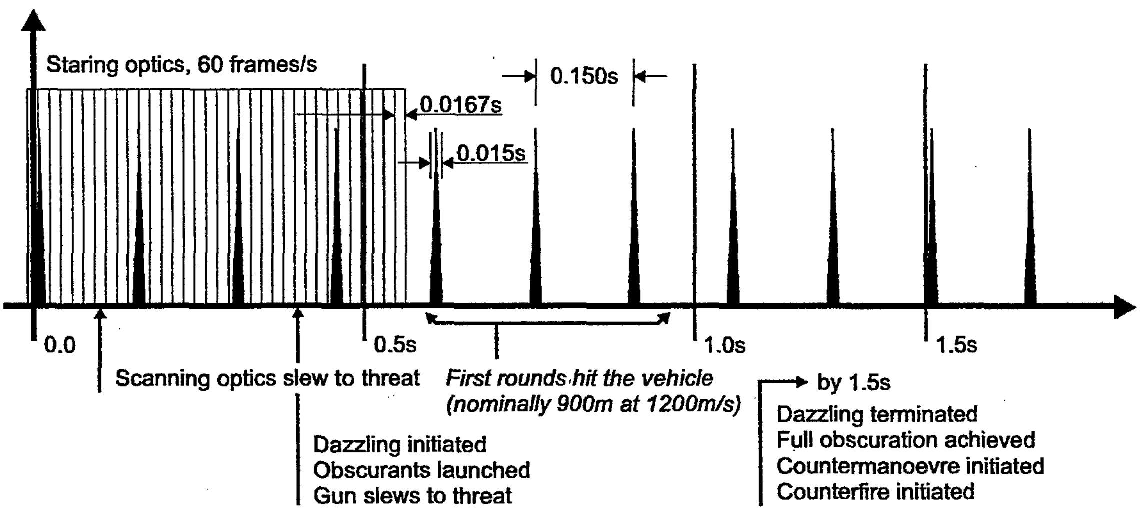
Typical grenade-burst pattern based on new LAV requirements, including a perimeter screen set at 40m, and for each quadrant three mid-level bursts at 45° and one at 70°. The original MBT screen for one quadrant is also shown.



For slowing, stopping and backing-up manoeuvres, a perimeter screen is set up with 4 grenades, a total of 5 mid-level grenades including 2 from aft launchers are used for additional coverage. For stationary vehicles, an additional grenade can be launched at 70 ° to counter sensor-fuzed submunitions.



Typical dazzling and grenade-burst patterns, automated for a moving vehicle. Five time intervals are shown. Dazzling is used to disrupt aiming or direct fire until the screen is in place.



An automatic weapon firing 400rds/min is detected by a staring array. A scanning optical system slews towards the threat and a dazzling laser is activated to disrupt the gunner. At the same time, smoke grenades are launched and the main turret slews towards the threat. By 1.5s, full obscuration is in place and the main gun can be fired using data from the Fire Control System or a Vehicle Network if available. These events are all stochastic in nature and can be analyzed in detail using ModSAF.

The author would like to thank Mr. Paul Brière and Dr. Gilles Roy for their useful advise and discussions on the subject of smoke grenade design and IR obscurants.

1. D. Sanschagrin, G. Couture, G. Roy and P. Brière, "Développement d'une Grenade Fumigène de 76 mm pour Véhicules Blindés: Correction du Mouvement Ascendant de l' Ecran Fumigène", DRDC-Valcartier Technical Memorandum TM-9510, February 1995.
2. G. Roy, P. Brière, D. Sanschagrin and G. Couture, "Development of a 76-mm Visual and IR Smoke Screening (VIRSS) Grenade for the Wegmann Launcher", DRDC-Valcartier Report R-9605, December 1996.
3. D. Sanschagrin, "Parameters Affecting the Performance of Immediate Screening Smoke for Armoured Fighting Vehicle Protection", DRDC-Valcartier Technical Memorandum TM-9829, December 1999.
4. D. Sanschagrin, "Effect of the Bulk Density of a Brass Flake Obscurant on the Dissemination of an Explosive Smoke Device (U)" DRDC-Valcartier Technical Memorandum TM-9836, December 1999.
5. D. Sanschagrin, P. Brière and J. Dumas, "Preliminary Field Trial of the Advanced Development Model of the 76-mm VIRSS Grenade (U)", DRDC-Valcartier Technical Memorandum TM-9837, December 1999.
6. J.L. Rapanotti, A. DeMontigny-LeBoeuf, M. Palmarini and A. Cantin, "Developing Defensive Aids Suite technology on a virtual battlefield", SPIE AeroSense Conference, April 2002.
7. R.G. Lee, T.K. Garland-Collins, D.E. Johnson, E. Archer, C. Sparkes, G.M. Moss and A.W. Mowat, "Guided Weapons", Bassey's (UK) Ltd., Third Edition 1998.
8. V. Larochelle, D. Bonnier, D. Dubé, G. Tardif and J. Bédard, "Optical Detection of Snipers Using Albedos: November 95 Trial (U)" DRDC-Valcartier Report R-9625, June 1997.
9. P. Webb, S. Soltesz, A. Cantin, J. Fortin and D. Pomerleau, "Improved miniaturized HARLID^{TM} for laser warning systems having high angular resolution", SPIE AeroSense Conference, April 2001.
10. J.L. Rapanotti, M. Palmarini and M. Dumont, "New computer and communications environments for light armoured vehicles", SPIE AeroSense Conference, April 2002.

Integrated Multispectral Camouflage for Mobile Weapon Systems (an effectiveness evaluation)

Bob Balma
Department of National Defence
Ottawa, Canada

ABSTRACT

The surveillance capabilities of armed forces throughout the world have increased tremendously during the past few years. The threat to military assets is multispectral and Camouflage, Concealment and Deception (CCD) measures must be provided in all the proper spectral regions in order to counter this threat. At the end of 1997, Director Soldier Systems Program Management (DSSPM) tasked Defence Research Establishment Valcartier (DREV) to determine the overall effectiveness of new mobile camouflage equipment against modern imaging systems from the ultraviolet (UV) to the thermal infrared (IR) spectral regions. An Integrated Multispectral CAmouflage for Vehicle Systems (IMCAVS) was designed by Barracuda Technologies of Sweden for the newly introduced Canadian Forces (CF) reconnaissance vehicle: the Coyote. To verify the enhanced characteristics of this new generation of camouflage equipment, a trial, under the umbrella of NATO, was conducted at CFB Valcartier in August 1998.

This paper presents results and comments on the specially designed concealment suite for the Coyote vehicle. This experimental concealment suite is designed to reduce the signature of the vehicle in the UV, visible and in both infrared spectral bands. It describes the design and characteristics of the Coyote concealment suite, a description of the experimental conditions and the instrumentation deployed during this trial. An indication of the results of a human perception experiment on the performance of the concealment suite in the visible band and electro-optical measurements taken in the UV and the two infrared (IR) bands are presented. Additionally, FLIR over flights conducted by CF188 aircraft are referenced. Finally, some conclusions and recommendations are tabled.

PROGRAMME OF WORK

The following milestones were established for the planned field trial:

a. Jan 98 – Signature analysis and design of camouflage suite;
b. Apr 98 – Plan of field trial;
c. Jun 98 – Production of hardware;
d. Aug 98 – Field trial;
e. Jan 99 – Review of data collected at field trial;
f. Oct 99 – National data evaluations (CA, GE, NL, IT, UK);
g. Mar 00 – Comparisons of national results; and,
h. Nov 00 – Final report to NATO RTO SCI (NATO CONFIDENTIAL).

TECHNICAL REQUIREMENTS FOR THE CAMOUFLAGE SUITE

A series of meetings were held between the Canadian Forces equipment users, the department of National Defence technical authorities and the camouflage producers to determine the requirements of the camouflage suite. The following requirements evolved:

a. to be designed for in-service combat vehicles;
b. to be for use in all tactical situations;

c. to be able to counter widely used sensors;
d. to be able to counter recognition and identification;
e. to be able to reduce the multispectral signature;
f. for use on moving combat vehicles;
g. enable rapid deployment and striking; and,
h. be a low cost add-on (not active or stealth)

APPLICATION OF SUITE WITH RESPECT TO COYOTE

The Coyote reconnaissance vehicle is a wheeled armoured vehicle that moves and stops frequently. It is fielded in two versions; one equipped with mast mounted sensors and the other with tripod mounted sensors having a remote capability. The stipulated technical requirements applicable to Coyote led to the development of a three stage camouflage suite (see Figures 1, 2, 3, and 4), as follows:

a. a permanent on board camouflage kit, (S1);
b. a permanent, onboard rapidly deployable short halt kit, (S2); and,
c. a long halt kit, (S3).

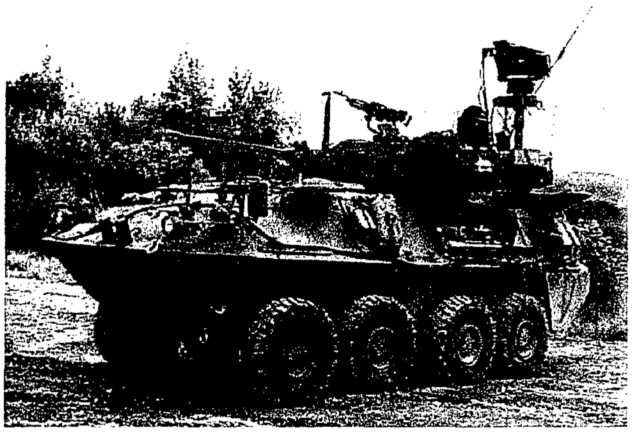


Figure 1 – The bare Coyote

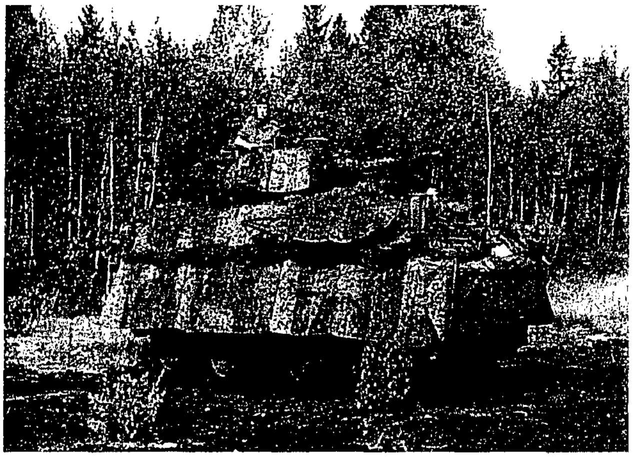


Figure 2 – The permanent camouflage (S1)

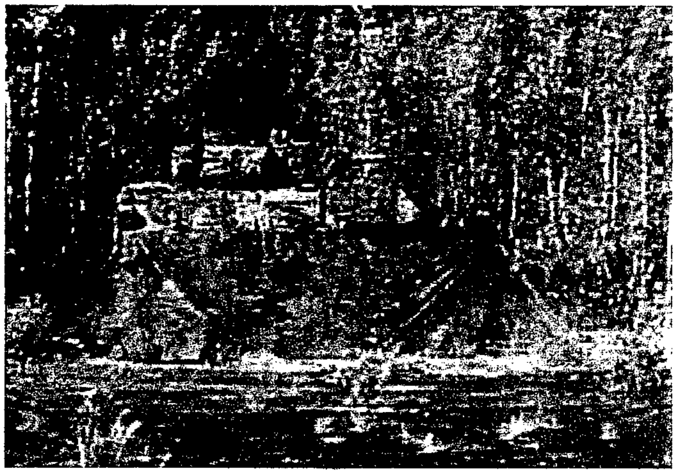


Figure 3 – The short halt kit (S2)

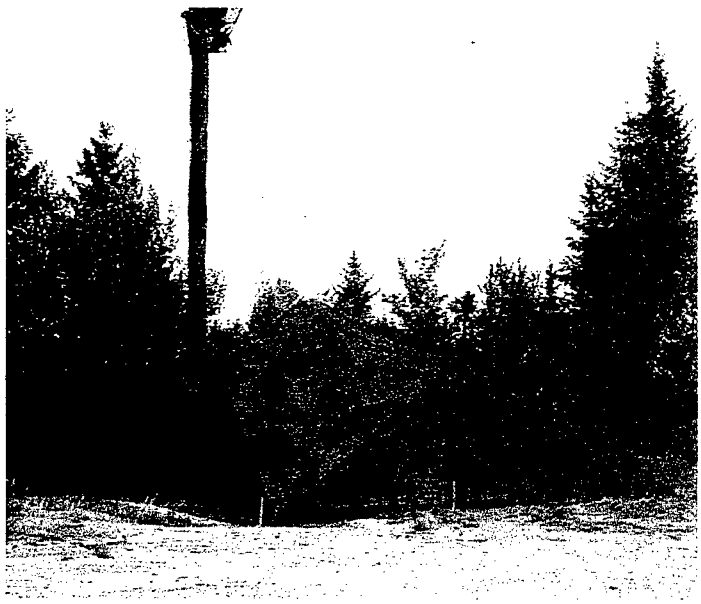


Figure 4 – The long halt kit (S3)

FIELD TRIAL BACKGROUNDS AND COYOTE DEPLOYMENT

The backgrounds selected into which the target vehicle was to be deployed were considered typical of the temperate northern hemisphere. The Coyote was deployed into what could be considered the most difficult of camouflage positions. As depicted in Fig. 4, one background was the front edge of a low wood line while the second background selected was in the middle of an open grassland plain interspersed with sandy patches.

DATA COLLECTION AND ANALYSIS

The data collections for the field trial consisted of:

a. imagery from helicopters;
   (1) 35 mm slides;
   (2) video (moving and still)
   (3) FLIR
b. calibrated imagery for ground truthing (UV, VIS, NIR, M &L Wave IR)
c. over flights from tactical aircraft with FLIR targeting systems

The analysis of the data sets was conducted by Canada, Germany, Italy the Netherlands and the United Kingdom. In several cases selected data sets were subjected to a round-robin analysis conducted in different countries. Several sets were analyzed using different models based on the same spectral domain.

TYPES OF ANALYSIS

The following figures (Figures 5 – 11) illustrate the types of analysis conducted and show unclassified comparative results:

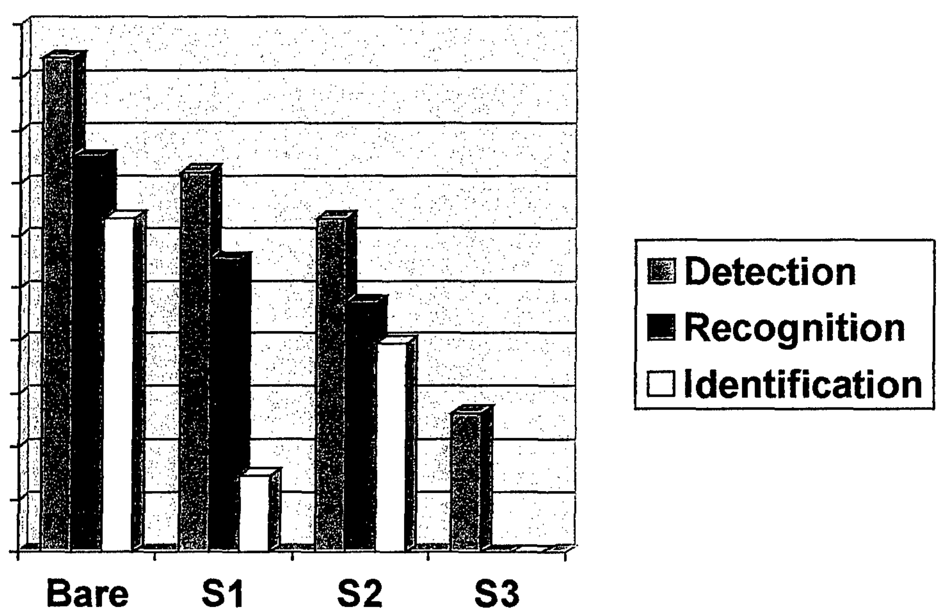


Figure 5 – Visual observation ranges in Metres

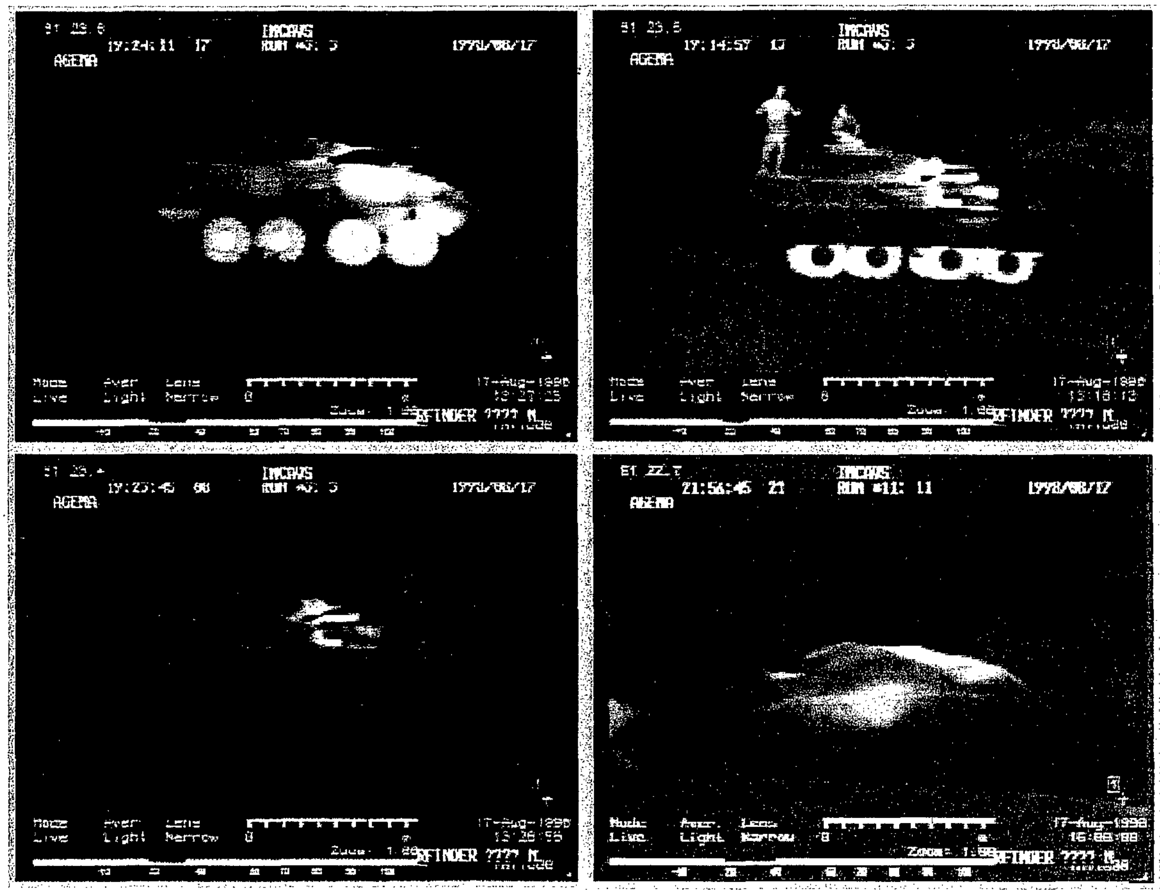


Figure 6 – LWIR ground truth imagery
(UL – Bare / UR – S1 / LL – S2 / LR – S3)

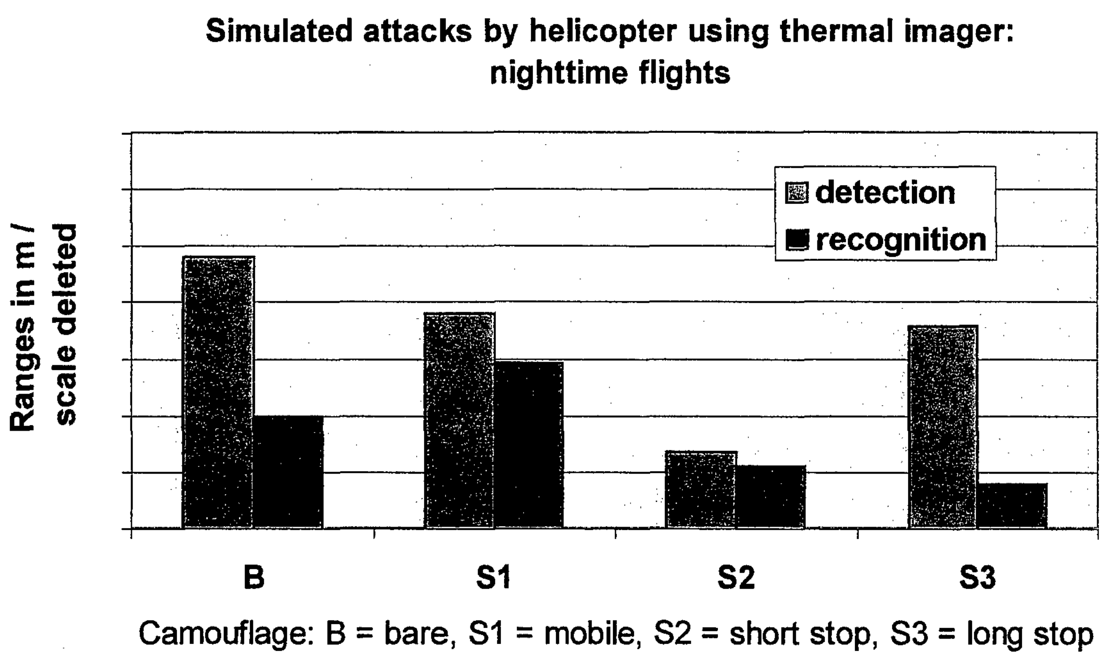


Figure 7 – Simulated attacks by FLIR equipped helicopter

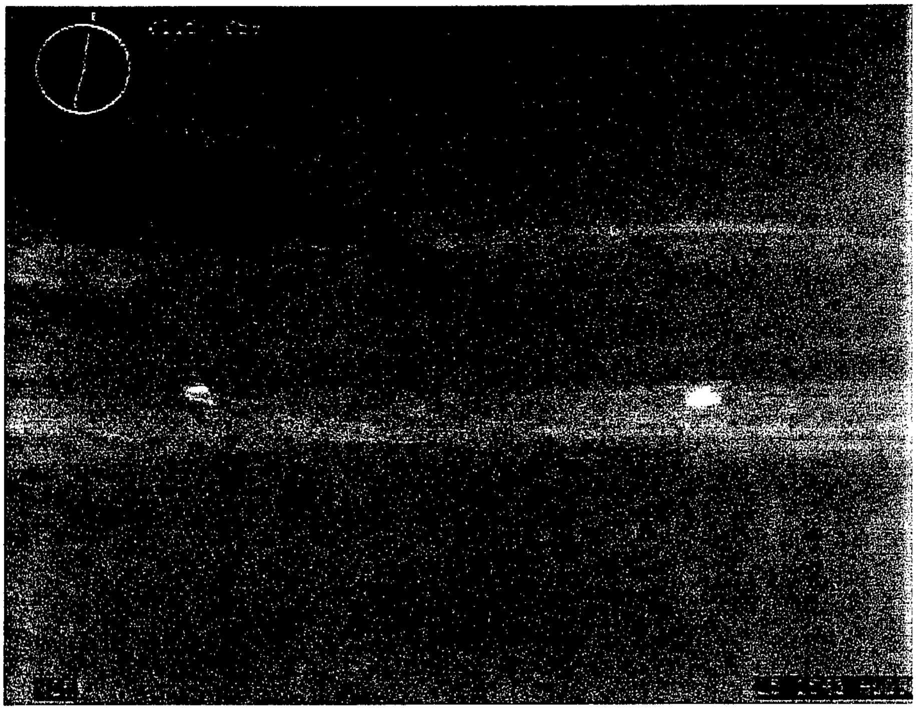


Figure 8 – Tactical MICROFLIR LWIR image of moving Coyote targets at 1 Km
RIGHT SIDE OF VEHICLES
(COYOTE LEFT - S1 / COYOTE RIGHT – Bare)

Figure 9 – Tactical MICROFLIR LWIR image of moving Coyote targets at 1Km
LEFT SIDE OF VEHICLES
(COYOTE LEFT – S1 / COYOTE RIGHT – Bare)

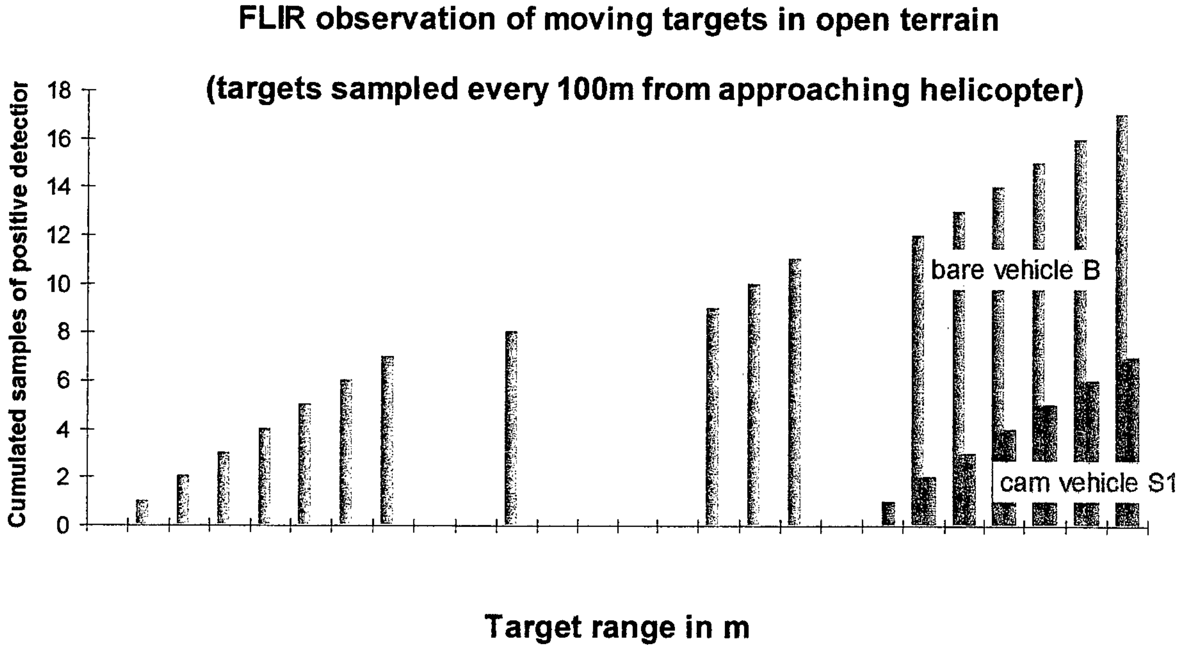


Figure 10 – FLIR observations from LONG to SHORT ranges (LONG ranges L)

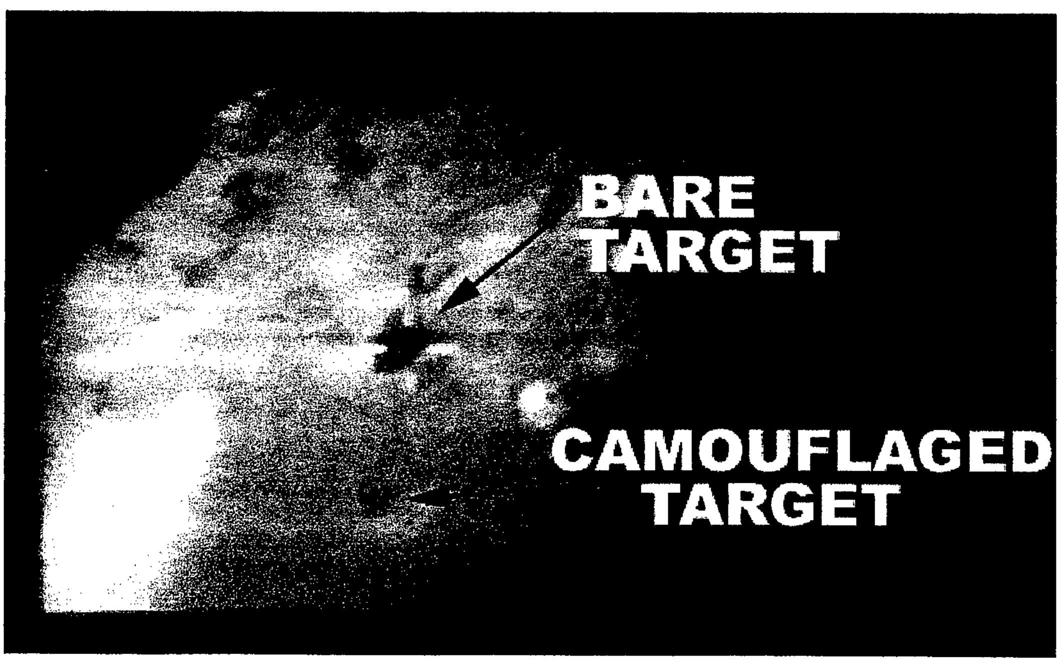


Figure 11 – Targeting FLIR image (BLACK hot) of Coyote targets
(engines running / 21:19 hrs local)

CUMULATIVE RESULTS OF THE ANALYSIS EXPRESSED AS CAMOUFLAGE GAIN

Figure 12 (below) represents the accumulated results of all of the analysis pertaining to the field trial. When used with judgment it is a useful tool for requirements staff in attempting to decide the "worth" of camouflage and signature reduction.

% REDUCTION IN DETECTION RANGE			
CONDITION	**PERMANENT**	**SHORT HALT**	**LONG HALT**
Day - Visual	20%	38%	66%
Night - Thermal	32%	72%	25%

Figure 12 – Accumulated results of field trial
(LONG HALT M&LWIR is a result of warm exhaust gases being trapped under the S3 kit)

SOME CONCLUSIONS FROM THE FIELD TRIAL

The design concept and engineering implementation of stages of camouflage proved beneficial in that:

a. various tactical situations could be addressed individually through the use of modules;
b. there was reduced workload for the soldiers;
c. a significant degree of multispectral signature reduction was achieved; and,
d. a basis for establishing the fiscal " worth" of signature reduction was achieved.

SOME RECOMMENDATIONS FROM THE FIELD TRIAL

The following general concealment considerations were documented:

a. Signature management must be addressed as an integral part of project management and be part of all initial engineering activities associated with new equipment procurement;
b. In-service vehicles considered to be high value assets should be retrofitted with customized camouflage kits appropriate to their role;
c. Vehicles that move rapidly and at short notice should be equipped with a suite of multispectral camouflage kits to facilitate concealment appropriate to their role and particular operational exigencies;
d. Means to direct and duct exhaust gases and to conceal equipment exhaust systems must be permanently installed;
e. The effectiveness of concealment suites must be tested in all climatic regions; and,
f. The lack of truly robust and cost efficient evaluation tools to measure camouflage effectiveness must be further considered.

OTHER RELATED NATIONAL PROJECTS

Figure 13 – QUICKCAM Kit for UK MBT CHALLENGER II

Figure 14 – Modular kit for GE MBT LEOPARD 2A4

REFERENCES

[1] Multiple authors, 2001, " Integrated Camouflage for Mobile Weapons Systems – An Effectiveness Study", NATO *RTO – TM 019AC/323-005)TP/30 NATO CONFIDENTIAL, Paris*, France.

Modeling the thermal signature of natural backgrounds

Marius Gamborg
Department of Electronics
Norwegian Defence Research Establishment
PO Box 25, NO-2027 Kjeller, Norway

ABSTRACT

Two measuring stations have been established - the purpose being to collect comprehensive databases of thermal signatures of background elements in addition to the prevailing meteorological conditions. The databases have primarily been used as a foundation for the development and validation of models for the simulation of thermal signature.

The stations consist of a calibrated thermal camera for radiometric measurements, as well as a number of meteorological sensors - i.e. sensors for recording relevant parameters that influence thermal signature. Both stations are remotely controlled via telephone lines and it is possible to transfer the collected data from the stations.

At each measuring site 3-4 thermal images are automatically recorded every 15 minutes and every 5 minutes over 50 meteorological measurements are performed. The stations have been operating successfully over long periods of time and they have delivered reliable meteorological information and radiometric data.

Using this data models for simulation of thermal signatures for natural backgrounds have been developed. In addition the measurements have been used for the verification of a model for vehicles. The background signature model is based on plain one-dimensional heat-flow equations. Applied to three different types of background elements, the RMS-deviations from measured temperatures appear to be less than 1.7 K for periods of several days.

INTRODUCTION

Previous work done at FFI has revealed a need to be able to correctly assess the effectiveness of thermal signature reduction measures [1]. The effect of these measures will depend on the thermal signature of the background. In order to be able to make such assessments it is necessary to know how the thermal signature of the background varies with factors such as type of terrain, weather, time of the year etc.

FFI-project 775 had as its purpose to establish comprehensive databases of meteorological data and thermal imagery, through the establishment of autonomous measuring stations [2, 3]. The databases were to be used for calibration and verification of two thermal models: One model for natural backgrounds, developed at FFI, and one model for vehicles. This paper describes the measuring stations, and the model for predicting thermal background signature.

MEASURING STATIONS

Measuring Sites

In order that the databases be as comprehensive as possible it was necessary with diverse meteorological conditions. Therefore three different measuring sites were chosen: Bardufoss, Ørland and Rygge. These sites are located in the north,

west and south east of Norway respectively. The geographical location and height above sea level of the three sites as well as their expected weather type is summarized in table 1.

Table 1: Measuring sites with expected type of weather

Measuring site	Location and height	Weather type
Bardufoss	69° 3'N, 18° 34'E, 100 m	Cold winter, polar night and short warm summer with midnight sun
Ørland	63° 43'N, 9° 38'E, 20 m	Wet and very windy winter
Rygge	59° 24'N, 10° 43'E, 55 m	Warm, dry summer

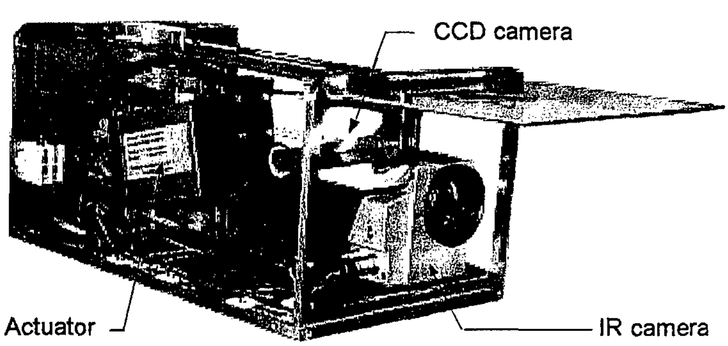
All the measuring sites are located at military air bases. This provides the necessary security to avoid tampering with the stations. It also provides a certain infrastructure, with electricity and phone lines available, even though the stations are situated close to uninhabited natural terrain. The air stations are located close to public weather stations, giving the opportunity to verify our own meteorological data against that of the public stations.

It was also important that a wide variety of background types were available at the sites, to be able to model several types of terrain. The sites chosen fulfil this need, containing coniferous and deciduous trees, scrub, heather, rocks, gravel roads, grass etc.

Camera system for radiometric measurements

The measuring stations consist of two separate systems. One system is a housing containing two cameras: An infrared camera and an optical camera. The other system consists of a data logger connected to several meteorological sensors.

The cameras are mounted together in a metal housing, which again is mounted on top of a 5-meter tall mast. The housing and the cameras are controlled automatically by a PC workstation. This is done with software developed at FFI. The housing is enclosed, and opens a door in front of the cameras only when images are recorded. When the door opens a fan blows heated air out through the opening to prevent any precipitation or moisture from entering the housing. The housing is also heated to prevent freezing inside. A picture of the camera system can be seen in figure 1.



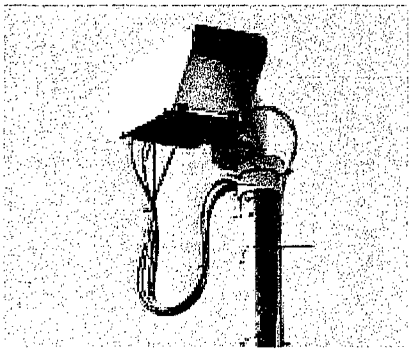


Figure 1: Camera housing without cover and with door open (left) and camera housing mounted on pan/tilt on top of the mast (right)

To be able to include as many different types of background elements as possible the field of view of the cameras is not fixed. The camera housing rests on a computer controlled Computar PT10 pan/tilt head. This pan/tilt can rotate and tilt the cameras, and thereby change their field of view. Operation of the pan/tilt is performed by the above-mentioned software.

Every 15 minutes this system records several visual and thermal images, and stores them on the hard drive of the controlling PC. The pan/tilt goes through a number of preset positions, and for each of them the cameras record images.

The distances from thermal camera to the scene is between 20 to 70 meters. At such short range the thermal transmission will not have a significant effect on the measurements except for during periods of dense precipitation.

Measuring Meteorological Parameters

The system for measuring meteorological parameters consists of several sensors connected to a Campbell Scientific CR10x data logger. Most of the sensors are placed among the background elements, and some have also been placed in a 5-meter tall mast. The mast is used for measuring wind direction and speed in addition to air temperature and humidity at two and five meter height. There are also four radiation sensors for measuring in- and out going radiation. A picture of the mast can be seen in Figure 2.

Sensors for measuring all parameters considered influential on thermal signature have been included: Air temperature and humidity, short wave (solar) irradiation, long wave (sky) irradiation, wind speed and direction, in addition to precipitation.

The logger takes measurements every 30 seconds. From calibration curves the data is then converted to normal units for temperature, wind speed etc. Every 5 minutes the logger stores an average of the last 10 measurements from each connected sensor in its internal memory.

Since the memory of the logger is restricted to data for only a few days it has been connected to the computer at the measuring site. It is possible to download the memory of the logger to the computer, thereby freeing logger memory.

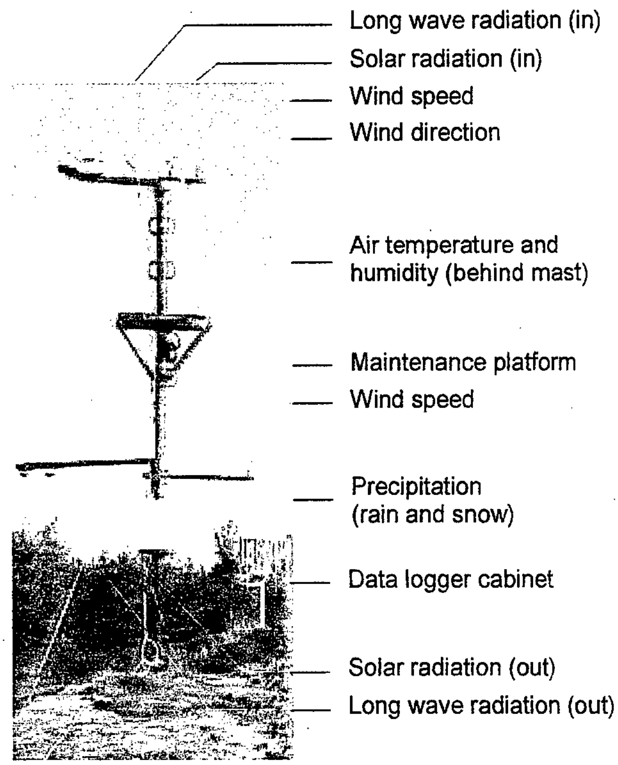


Figure 2: Mast with meteorological instruments. Precipitation sensor in background

Measuring Thermal Signatures

The measurement of thermal signature is based on the thermal images from the infrared camera. Two different types of cameras are used: FLIR Instruments ThermaCAM PM 595 and Agema Systems Thermovision 570. The two types have the exact same specifications: A spectral range from 7.5 to 13 μm, a field of view of 24° x 18°. Both cameras give images of 12-bit depth with a resolution of 320 x 240 pixels. The cameras are calibrated and have a thermal sensitivity of 0.1 K.

By averaging the pixels covering one background element the apparent temperature of that element can be found. Repeating this operation for several consecutive images gives the temperature variation over time. An illustration of this process is shown in figure 3.

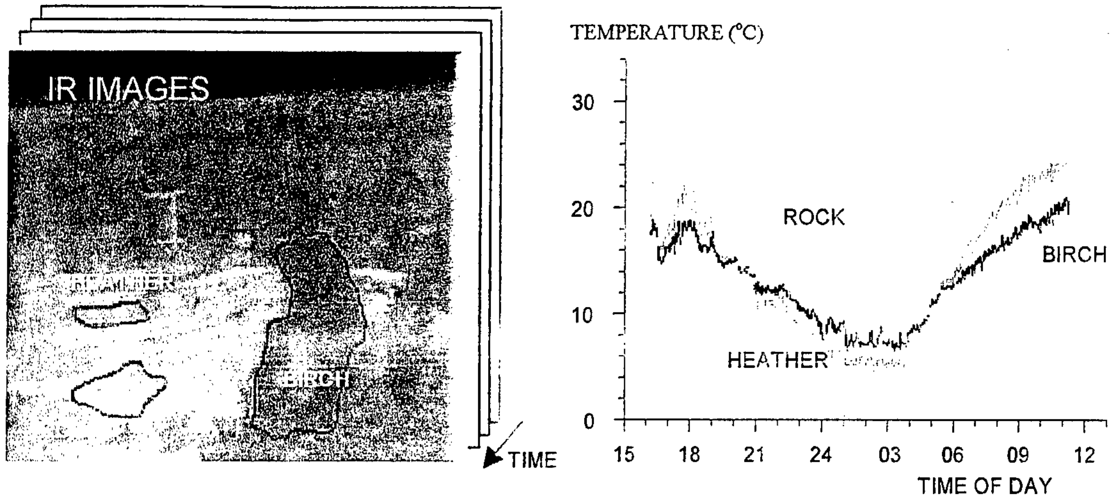


Figure 3: IR images used for studying temperature variation over time. The curves on the right are average temperatures from the background elements plotted versus time.

BACKGROUND MODELING

There have been developed several mathematical models to simulate thermal signature. Several of these models utilize very complex physical descriptions of the problem. Hence they require high computing power, advanced mathematical solutions and may require a long time to converge on a solution.

On FFI-project 775 it was decided to go with a simple solution due to the short time available. It was seen as desirable to limit the computing power needed for simulations to single workstations. A very promising model has been developed at FGAN-FOM, and the choice of model was based on the results from this model [4].

Description of the General Background Signature Model

The signature model presented in this paper is based on heat transfer from objects with only one surface exposed to the surroundings. To further simplify, it is modeled as a one-dimensional heat transfer. The background elements are divided into uniform layers, with only the top layer being exposed to the environment. All the other layers simply transfer heat through one-dimensional conduction. The mathematical equations that describe this process are presented in equation 1.

$$
\begin{aligned}
\frac{dT_s}{dt} &= \alpha \cdot W_{sun} + \varepsilon \cdot W_{sky} - \varepsilon \cdot \sigma \cdot T_S^4 && \text{Radiation} \\
&\quad - (\delta_1 + \delta_2 \cdot v_a) \cdot (T_s - T_a) && \text{Convection} \\
&\quad + \rho \cdot (\delta_1 + \delta_2 \cdot v_a) \cdot (e(T_a, Rh) - e(T_s, 100\%)) && \text{Evaporation} \\
&\quad + \kappa \cdot (T_1 - T_s) && \text{Conduction} \\
&\;\cdots\cdots && \qquad (1) \\
\frac{dT_i}{dt} &= \kappa \cdot (T_{i+1} - 2 \cdot T_i + T_{i-1}) && \text{Conduction in layer no. i} \\
&\;\cdots\cdots \\
\frac{dT_n}{dt} &= \kappa \cdot (T_c - 2 \cdot T_n + T_{n-1}) && \text{Conduction in layer no. n}
\end{aligned}
$$

In equation 1 α, ε, σ, ρ, κ, δ_1 and δ_2 are the parameters of the model, T_s is the surface temperature of the background element, T_a is the air temperature and T_i is the temperature of layer number i. T_c is the temperature of the lowest layer, and does not vary with time. The meteorological data required is W_{sun}, W_{sky}, v_a, T_a and Rh, which are the short and long wave irradiation, wind speed, air temperature and humidity respectively. e (T_a, Rh) is the water vapour pressure given by the air temperature and humidity.

When the values for the parameters are known, the surface temperatures are calculated by converting the differential equations in equation 1 into difference equation of a certain time step. For each time step the temperature change is calculated and added to the temperature of the previous time step. The input required for this operation is a starting temperature profile, i.e. one temperature for each layer, as well as meteorological data for each time step. Thus a curve of surface temperature versus time is generated, based on certain values of the model parameters.

The fit of the model is calculated as the root mean square error, or the RMS error, between the model curve and the temperature for the relevant background element as extracted from the thermal images. Varying the parameter values until the RMS error is minimized results in the final estimation of the parameters.

Minimizing the RMS error is a trivial problem for linear functions. However this is not generally the case for differential equations. We have therefore used a combination of the simplex "downhill method", and the Levenberg-Marquardt method [5, 6].

General Edge-of-Forest model

This model is a general edge-of-forest model. It has been developed on the basis of both coniferous and deciduous trees and is therefore valid for most types of trees. However the model is less complex than the general model presented above.

In this model there is no conduction, as this factor will have a negligible effect on trees. In addition to this, the humidity measurements did not seem to have significant impact on the results, and therefore the humidity term was dropped to simplify further.

A plot of the edge-of-forest model together with actual measured temperatures as well as a best-fit curve appears in figure 4. The best-fit curve is a plot of the parameter values that gave the smallest RMS error for the days plotted. It gives very good results for the period shown, but does not perform as well as the general edge-of-forest model for other periods.

The general edge-of-forest model was developed on the basis of several 5-day sequences from a 9-month period. It has been plotted for the entire period of nine months. For this period it gave an RMS error of about 1.9 K. It is however not that difficult to get these results for the edge-of-forest, since foliage temperature has a tendency to follow that of the air quite closely.

The average difference between the edge-of-forest model and the measured temperatures has been plotted in figure 5. As can be seen from the plot the model accuracy degenerates rapidly below about 260 K. This can be explained by several factors. As the temperature drops below freezing, icing may affect the reflective properties of the trees. Also the phase changes back and forth between ice and water will absorb and liberate energy. In addition the fact that low temperatures will occur mostly during periods when there are no leaves on the deciduous trees will change their thermal properties. Finally there are fewer data containing low temperatures, meaning that the parameter estimation will weigh the lower temperatures less.

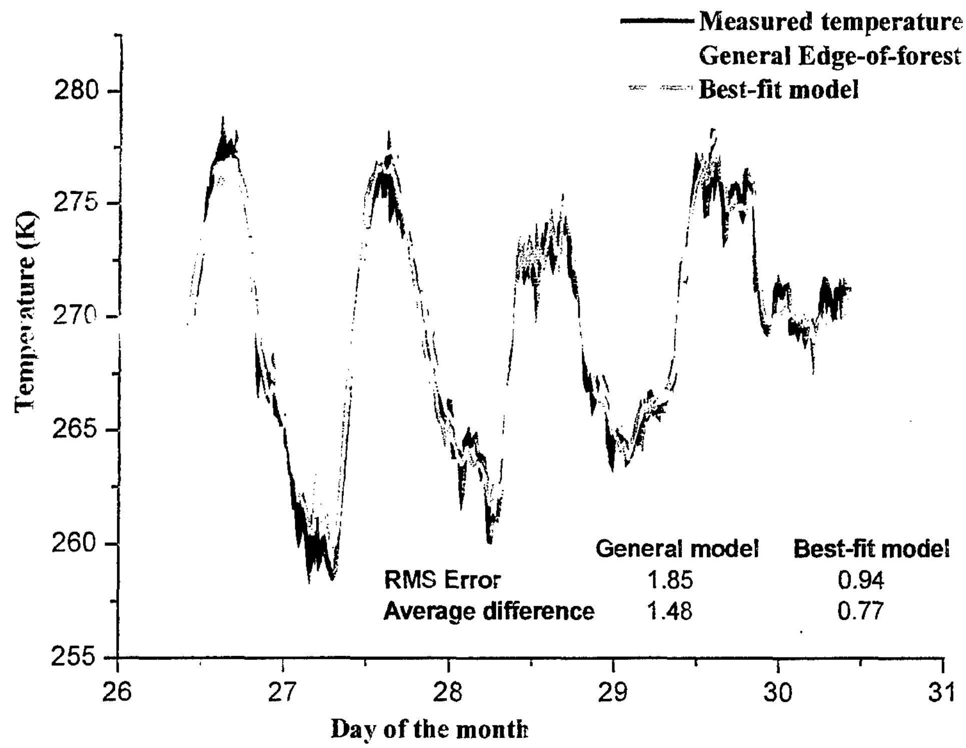


Figure 4: Plot of Edge-of-forest models for late March 2001, with values for average difference and RMS error inserted.

Rock model

Like the edge-of-forest model this model is a simplification of equation 1, since also here the air humidity has been excluded. Modeling of the rock has been tried both with and without the conduction terms. Without the conduction terms the model performs quite well for short periods of time. However, it does not perform as well when extended beyond a few days.

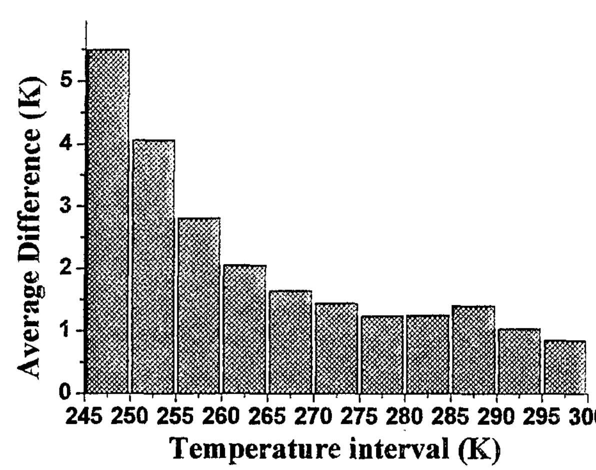


Figure 5: Average difference between general edge-of-forest model temperature and actual measurements

The conduction is modeled by dividing the rock into 20 layers vertically. The conduction coefficient is assumed to be equal between each layer. No assumptions are made about the thickness of the layers, other than that the lowest layer will be sufficiently deep to be at a constant temperature. This constant temperature was set to 277 K. The initial temperature of the 20 layers is simply a linear interpolation between the measured surface temperature and the constant temperature of the lowest layer. This means that the model will need some time to adjust, since the initial ground profile does not contain any temperature history. The conduction is assumed to be constant with depth, and independent of the time of year, wetness etc.

A plot of the rock model for 5 days in May 2001 is plotted in figure 6. The plot contains a best-fit curve and a curve for a long-term model. The best-fit curve has been fitted specifically for the last part of May, while the long-term model has been verified for the period of early March to late June. As can be seen from the plot the long-term model has problems with the highest temperatures. This is probably due to the fact that these are extreme temperatures for the period upon which the model is based. The RMS error of the long-term model is 2.3 K for the entire three-month period. The segment of the long-term model presented in figure 6 represents the period with highest RMS error.

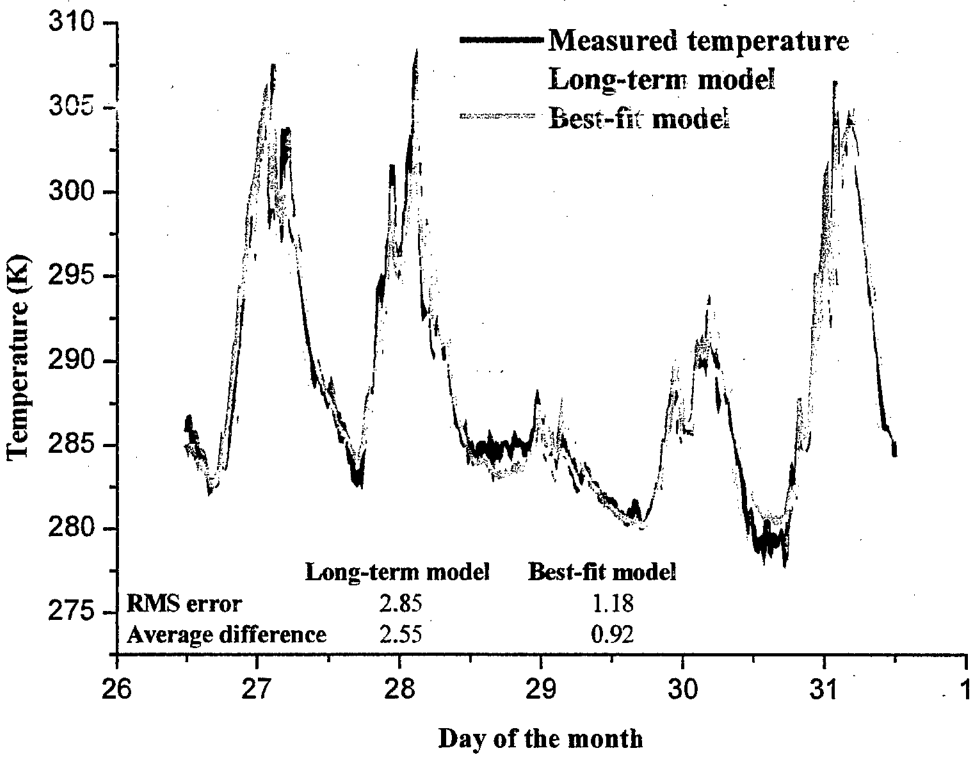


Figure 6: Plot of rock models for late May 2001, with values for average difference and RMS inserted.

Model parameters

Model parameters for the general edge-of-forest model and for the long-term rock model are presented in table 2. The model values are the parameter values that are input into the model. The modified values are the model values multiplied by a factor so that the long wave emissivity coefficient ε becomes 0.95. The value 0.95 was chosen as a fairly common value for the long wave emissivity coefficient for natural materials. The multiplication factor represents the heat capacity of the elements modeled. The only parameter that has not been modified by the heat capacity is σ, as this term only appears in the model in conjunction with ε.

As can be seen from table 2 the parameters take values that do not seem "unphysical" in nature. There are no negative parameters, which would imply cooling of the background element when energy is added to it. The short wave absorptivity coefficient α is between 0 and 1, but deviates slightly from tabulated values [7]. In addition Stephan-Boltzmanns constant σ is close to its actual value of 5.67e-08 $Wm^{-2}K^{-4}$. The reason for the deviations is probably that the model to a certain extent accounts for effects not incorporated in it.

FURTHER WORK

The models need to incorporate effects that have not been considered in the first version. This includes solar angle of incidence for elements that are not horizontal, precipitation and possibly thermal irradiation from surrounding terrain.

Table 2: Parameter values for the general edge-of-forest model and for the long-term rock model

	Edge-of-forest model		Rock model	
Parameter	**Model Value**	**Modified value**	**Model Value**	**Modified value**
α	2.66e-06	0.25	5.92e-06	0.63
ε	1.00e-05	0.95	8.89e-06	0.95
δ_1	1.40e-04	13.30	9.06e-06	8.25
δ_2	7.00e-05	6.65	7.72e-05	0.97
σ	5.84e-08	5.84e-08	6.40e-08	6.40e-08
κ	N/A	N/A	1.22e-05	12.99

There are also plans to model snow. In this case it will also be necessary to record the condition of the snow in some way. There are no plans to incorporate the physical three-dimensional structure of the background elements, as this will make the model unnecessarily complex.

If time allows it there are plans to make one of the measuring stations mobile. This will allow for measurements in other types of climates than only in the three fixed measuring sites chosen so far. However, this operation will be very time consuming.

Currently the measuring stations are only used for purposes of delivering data for the model calibration and evaluation. However they are ideal for evaluating vehicle and personnel thermal camouflage. The stations will to a certain extent be made available for the Norwegian military to use for this purpose.

Finally, the models and the collected data will be used as an aid in assessing the effectiveness of thermal sensors. When parameters have been estimated for all kinds of backgrounds it will be possible to estimate the temperature variation in given backgrounds for given meteorological conditions. If this is used in conjunction with calculations of atmospheric propagation the performance of thermal sensors can be evaluated.

CONCLUSION

The measuring stations can deliver data that is sufficiently accurate to be used for modeling of thermal signature. The use of a camera instead of a radiometer has allowed a more precise pinpointing of the location of relevant background elements, thereby increasing the accuracy of the measurements. It also gives the possibility of studying the temperature variations over the background element.

So far two general sets of model parameters have been estimated: One model for a general edge-of-forest, and one for bare rock. The edge-of-forest model has been plotted for a period of nine months, giving an RMS error of 1.9 K. The rock model has been plotted over a period of 3 months, resulting in an RMS of 2.3 K.

The models have resulted in parameters that seem reasonable from a physical perspective. Some of the parameters deviate slightly from tabulated values, but this is not unexpected as the model is much less complex than the thermal systems it simulates.

Modeling has shown that even with very few parameters it is possible to obtain model parameters that result in a very good model for periods of at least two months. Such results can be obtained even for parameters that obviously represent

"unphysical" values, due to the existence of local minima in the fit-function. Therefore it is important when modeling to try different starting values for the parameters, as the mathematical methods used for estimating the parameters may not always converge on the global minimum.

The models developed are simple and quick. They can be used to estimate thermal signatures in short time on a single workstation. The accuracy of the models is sufficient to apply to the evaluation and assessment of thermal camouflage, and will therefore be a valuable tool in further work at FFI.

AKNOWLEDGEMENTS

This paper presents work done at FFI-project 775. I would like to extend my gratitude to my co-workers on the project: F B Olsen, E Strømman, T Høimyr and M Søderblom.

The success of this project has depended on the continuous operations of the measuring stations. Even though they are autonomous there have been several problems along the way. To keep the stations in operation despite these problems it was necessary to have someone to look after them. Therefore one person at each of the air stations was appointed as contact person. The cooperation of the contact persons J E Sjøgren, Major S A Karlson and Major P K Wikestad at Rygge, Ørland and Bardufoss respectively, is very much appreciated.

Thanks also to FGAN-FOM, for providing all the equipment for the measuring station at Bardufoss, and for giving us valuable information on the development of their own thermal signature model.

REFERENCES

[1] Skjervold J E, Strømman E, Olsen F B, Høimyr T (1998): Operational inclusion of multi spectral camouflage in the Army – Final report for FFI-project 672, *FFI/RAPPORT-98/05956*, Kjeller, Restricted (In Norwegian)

[2] OLSEN F B, Gamborg M. Høimyr M, Strømman E, Søderblom M (2001): Thermal signatures of objects and backgrounds - Final report for FFI-project 775, *FFI/RAPPORT-2001/05210*, Kjeller (In Norwegian)

[3] Strømman E, Høimyr T, Gamborg M, Olsen, F B (2001): An automatic station for measurement of meteorological parameters and thermal signatures, *FFI/RAPPORT-2001/05402*, Kjeller

[4] D. Clement, W. Jessen (1993): A Background Model in the Thermal Infrared: Status, Validation, and Applications, *FGAN-FfO*, 1993/41

[5] Marco S. Cacceci, William P. Cacheris (1984): Fitting Curves to Data. The Simplex algorithm is the answer, *BYTE*, May 1984, 340-362.

[6] William H. Press et al (1992): Numerical recipes in C: the art of scientific computing, *Cambridge University Press*, Cambridge, 408-412, 683-685.

[7] J.L. Monteith, M.H. Unsworth (1990): Principles of Environmental Physics, *Edward Arnoldl*, London, 80-82.

[8] Gamborg M, Olsen F B (2001): A database for thermal background signatures and their corresponding meteorological parameters - Data organization and reduction, *FFI/RAPPORT-2001/05482,* Kjeller

[9] Olsen F B, Gamborg M (2001): Modelling the thermal signature of natural backgrounds, *FFI/RAPPORT-2001/05324*, Kjeller

High-resolution Ground Target Infrared Signature Modeling
For Combat Target Identification Training

Jeffrey S. Sanders
Signature Research, Inc.
150 West Park Loop, Suite 203
Huntsville, AL 35806
(256)721-9009

Abstract

Recent world events have accelerated the evolution of the US military from monolithic formations arrayed against a known enemy, to a force that must respond to rapidly changing world events. New technologies are part of the Army's evolution and thermal imaging sensors are becoming more and more prevalent on the modern battlefield. These sensors are integrated into advanced weapon systems or commonly used for battlefield surveillance. Thermal imaging systems give the soldier the ability to deliver deadly force onto an enemy at long ranges at any time of day or night. The ability to differentiate friendly and threat forces in this situation is critical for the avoidance of friendly fire incidents and for the proper use of battlefield resources. The ability to foresee the location of the Army's next battlefield is becoming more difficult, and we don't know where the next battlefield will be from year to year. Infrared target recognition training tools need to be flexible, adaptable, and be based on not only the latest intelligence data but have geographically specific training available to the soldier.

To address this training issue, personnel of the Measurement and Signatures Division at the National Ground Intelligence Center have created the Simulated Infrared Earth Environment Lab (SIREEL) web site. The SIREEL web site contains extensive infrared signature data on numerous threat and friendly vehicles and the site is designed to provide country-specific vehicle identification training in support of US military deployments. The bulk of the content currently on the site consists of infrared signature data collected over a decade of intelligence gathering. The site also employs state of the art infrared signature modeling capabilities to provide the soldier in training the most flexible training possible. If measured data on a vehicle is not available, the website developers have the capability to calculate the infrared signature of ground vehicles in any location, any type of terrain, any weather condition, any operational state, at any time of day on any day of the year. This allows the SIREEL website developers to completely populate target signature training databases when measured data is unavailable for required vehicles. This paper explores the methodologies and tools necessary to provide the predictive infrared ground vehicle signatures for this application.

Introduction

When developing a concept for training soldiers how to interpret the output of an infrared sensor there are multiple critical issues to consider:

1) The training system must be able to provide information on any possible target that the soldier will encounter during a given deployment.
2) The training system should be able to be tailored to a particular deployment so that the soldier in training is not overwhelmed with learning requirements.
3) The signature data used to train the soldier must be of high quality and contain all meaningful feature information required to recognize a particular vehicle.
4) The vehicle signature data must be able to be presented as a part of a realistic infrared scene that is representative of the region of an impending deployment.

The four issues mentioned above define some of the basic requirements for an infrared signature combat identification training system. The first issue requires that the training system have a complete inventory of vehicles that a soldier can be expected to encounter on a given deployment. Gaps in the training system could lead to situations on the battlefield where a soldier encounters a target that is unknown to him and this could lead to delayed reaction times and increased risk to personnel.

The second major issue above, which relates to deployment-specific training, is important because the soldier should not have to attempt to retain information that is not relevant to his current mission. If he spends valuable training time memorizing signatures of vehicles that will not be encountered, his effectiveness is reduced and he is unnecessarily burdened with useless information. The signature of a vehicle in its original country of manufacture can be changed drastically due to the complex proliferation of military vehicles throughout the third world due to country-specific modifications. The soldier should have available the signatures of the vehicles as they appear in a particular country or region. For example, the Iraqi Type 69-II main battle tank (MBT) is a heavily modified Iraqi variant of a Chinese copy of a Russian T-55. While in many respects it is a T-55, its outward appearance is drastically different from the standard Russian MBT. To be effective, the training system should contain signature data on all the relevant variants and configurations of a given base vehicle.

The third training requirement issue is a basic quality requirement in that the signature data being used to present training information should be high quality signature data and be able to provide all pertinent IR feature information to the soldier in training. The data should represent as many operational states and environmental conditions as necessary. When a soldier views a target through an infrared sensor there are an infinite number of possible ways that a given target can appear. The target can have country-specific modifications and could possibly have some type of passive countermeasure such as camouflage employed. How the vehicle appears is also a strong function of the natural environment it is in, its operational state, the time of day, and the time of year. An infrared combat identification training system should be able to provide all of these signature states as needed for training.

The final general requirement for an infrared combat identification training system is the ability to present vehicle signatures imbedded in realistic infrared scenes representative of a specific region of interest. The vehicle detection and identification process is strongly dependent upon not only the signature of the vehicle itself but the scene clutter in which the vehicle resides. In addition, in many combat situations a target vehicle could be in defilade or partially occluded by objects in the scene. The ability to present partial signatures of vehicles is an important capability for realistic training.

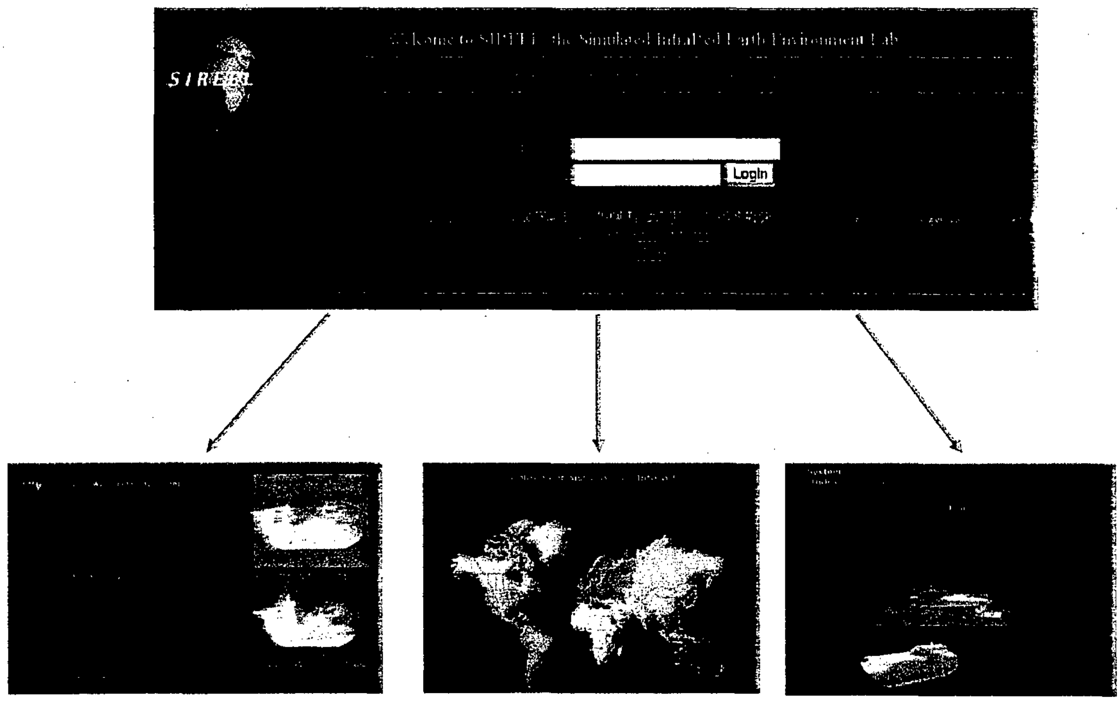
The infrared combat identification training requirements mentioned so far carry with them the implication that any training system must have an incredibly large number of signatures available for adequate training. This is driven by the large number of military vehicles presently in existence that could be encountered on a battlefield, the different signature states of each vehicle, the variants of each vehicle, and a given vehicle's configuration. While using measured signatures is the obvious place to start with a training system, it is simply not possible to collect signature data for every desired training scenario. Fortunately, there have been recent technological advances in the field of ground vehicle infrared signature prediction and computer processing power so that it is now feasible to address these signature requirements with predictive infrared signature models.

First principles of physics predictive infrared signature modeling provides the capability to calculate the infrared signature of a vehicle in any operational state, at any time of day or year, at any location in the world, and under the influence of any type of weather. Once a high-resolution predictive signature model has been built it can be used to generate as many signatures as required for training. These predicted signatures can be validated against measured signatures and there are numerous examples over the history of this field where it has been shown that it is possible to create a model that matches measured signatures very closely. Validated, computer-generated signatures offer a huge long-term cost savings over measured signatures. Currently on the SIREEL program it is possible to generate ground vehicle infrared signatures in volume at a cost in the tens of dollars per signature with a generation time measured in days. On the other hand, measured signatures could easily cost thousands of dollars per signature when the costs of vehicle operation, range time, and sensor costs are added up. A limited number of these relatively expensive measured signatures will be required for validation of the predictive models but once they have been validated to the required level of confidence, predictive models offer substantial cost savings.

SIREEL Website Overview

The SIREEL website consists of two major components, a general IR signature training section and a deployment section that provides country-specific information. The site is intended to provide simple, basic instruction on IR signatures while having an extensive and comprehensive database of vehicle signatures for the user to learn from as needed. A unique feature of SIREEL is the country-specific information that allows a user to learn about vehicles that will be encountered on any given military deployment. Figure 1 shows a conceptual layout of the SIREEL website.

SIREEL Secure Server Login Page



Infrared Signature Basic Education | **Country-specific Deployment Data** | **Extensive Vehicle Index**

Figure 1. SIREEL website contents and layout

The SIREEL site is currently accessible over the Internet and requires that a potential user acquire a login and password from the website administrator. The site is a secure site requiring 128-bit encryption but is unclassified and can be accessed from any military or government Internet domain. The SIREEL site is constantly being improved and populated with additional infrared signature data of ground vehicles, helicopters, and aircraft. Hundreds of vehicles must be represented to provide a comprehensive training set for target identification training with the primary focus being on ground vehicles such as tanks, infantry fighting vehicles, rocket launchers, and other primary battlefield targets. As mentioned previously, this requirement for signatures of exotic targets in locations all over the world necessitates the use of predictive IR signature modeling techniques that are based on first principles of physics.

SIREEL Infrared Target Signature Modeling

The use of predictive infrared signature models on the SIREEL website allows for the presentation of representative vehicle signatures for a vehicle at any time of day, in any weather condition, anywhere in the world. This capability is required in order to create geographically specific IR scenes for training. While measured IR signature data is very useful for demonstrating fundamental signature features of a vehicle, it is impossible to collect enough data to represent a vehicle's IR signature at every point in the vehicle's

multi-dimensional signature space. This necessitates the use of predictive IR signature models.

The desired result of the ground vehicle IR signature modeling process is to end up with a 3D model that can be manipulated over the web for training purposes. The model must be able to convey important information such as vehicle shape from different aspects and location of thermal features. The process for creating an IR web model of a vehicle consists of four major steps: creating a geometric representation of the vehicle, deriving a thermal mesh from the geometric reference, attributing the thermal mesh and creating a predictive signature model, and creating an optimized real-time model compatible with web browser visualization software. Figure 2 shows these steps and examples of models at each stage of the development process.

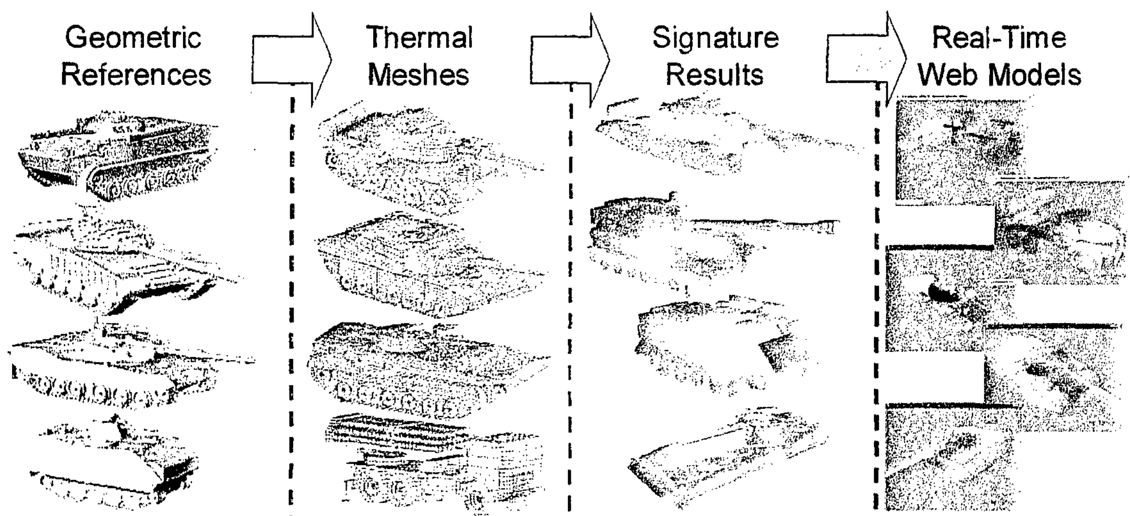


Figure 2. The Four Stages Required for IR Signature Web Models

The ground vehicle IR signature modeling process begins with the creation of a three dimensional geometric reference model of a vehicle. This has historically been a difficult task, but the speed and capabilities of personal computers, and the capabilities of modeling software have increased to the point where this stage of the development can be accomplished with commonly available tools. The most difficult challenge typically faced by a geometry modeler is building a model that is sufficiently accurate for a given application. Creating models for target training applications often involves vehicles about which very little is known. In some cases assumptions must be made concerning the shape and size of components of the vehicle because the best information available may be photographs. In other cases, accurate line drawings or examples of the actual vehicle may be available and the accuracy of the model can be very high. Fortunately, IR signature models have less stringent spatial detail requirements than other types of signature models and IR signature models can be created with sufficient detail from relatively limited information. The typical spatial resolution of the SIREEL predictive signature models ranges from three to five inches and non-critical features smaller than this are not modeled.

After the geometric reference has been created it is used as the primary inputs for the creation of a thermal mesh. The thermal mesh is simply another representation of the geometry of a vehicle but with very specific requirements. The thermal mesh must be created with low aspect ratio quadrilateral polygons in order to properly model the flow of heat across a vehicle component. In the final thermal solution, each polygon is assigned a calculated temperature over time based on a number of heat inputs over time, and each polygon is referred to as an isothermal node. The thermal signature codes have a finite limitation on the number of node temperatures that can be calculated at one time and this limitation drives the final resolution limits of the signature model. Figure 3 shows a typical thermal mesh created for the SIREEL website. The mesh shown in figure 3 is comprised of approximately 45,000 quadrilateral or triangular polygons.

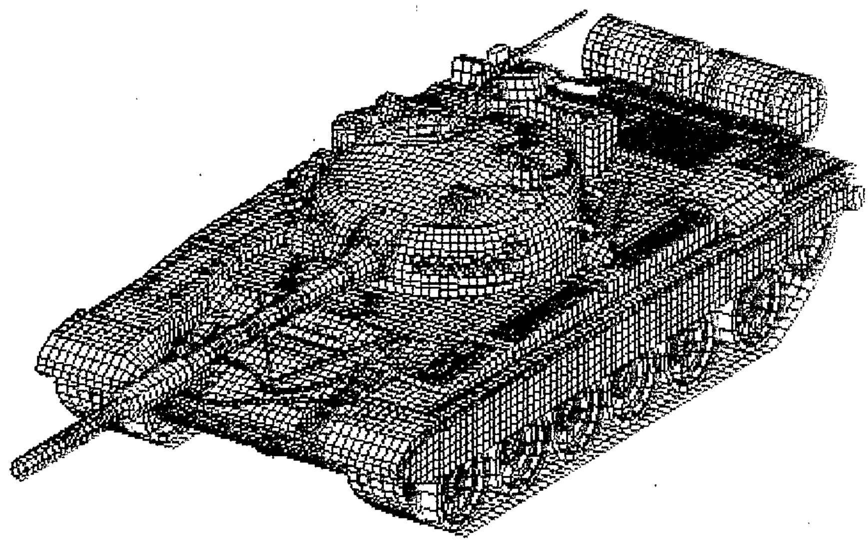


Figure 3. Typical Ground Vehicle Thermal Mesh

After the thermal mesh has been generated and segmented into appropriate groups, the mesh is used as the primary input to a signature prediction code. The currently available choices for ground vehicle signature prediction codes are the Physically Reasonable Infrared Signature Model (PRISM), and the Multi-Service Electro-optic Signature (MuSES) code. PRISM is a code that has been the historical standard for ground vehicle signature prediction and MuSES is the recent intended replacement for PRISM. MuSES is the more modern code with a more user friendly interface but it is a proprietary code and the user does not have access to the source code. This lack of access to source code makes it difficult for the user to model complex heat transfer processes such as the ones occurring in vehicle engines. PRISM source code is available and this allows the PRISM user to create whatever functions and subroutines necessary to model whatever is required. However, due to its more modern software architecture MuSES can accommodate models with higher numbers of thermal nodes and this allows for higher spatial resolution models. The approach used for the SIREEL infrared signature models is a hybrid process that takes advantage of the strengths of both codes. A custom version of PRISM was created that can accommodate more nodes than the standard release version of PRISM, and for relatively low node count models, PRISM was used to

generate the IR signature. For models with higher node counts, PRISM was used to calculate the temperatures of engine components and temperature curve files were written out for importing into MuSES. MuSES calculates environmental interactions, node to node conduction and node to node radiation exchange. These fundamental heat transfer calculations combined with the engine component temperature curves from PRISM represents a highly flexible hybrid approach to ground target IR signature generation.

The engine, drive train, and running gear of a ground vehicle are the dominant contributors to a vehicle's IR signature in most tactical scenarios. The heat generated by the engine and drive train propagates from these active components to other parts of the vehicle by complex radiation, conduction, and convection paths. The temperatures of the components of the running gear (wheels and tracks) of a ground vehicle are driven by parameters such as the speed of the vehicle, mass of the vehicle, terrain type, and bearing friction. All of these complex heat transfer mechanisms must be addressed in order to accurately calculate the IR signature of an active ground vehicle. Figure 4 demonstrates the general process for the incorporation of detailed engine models into ground vehicle signature models.

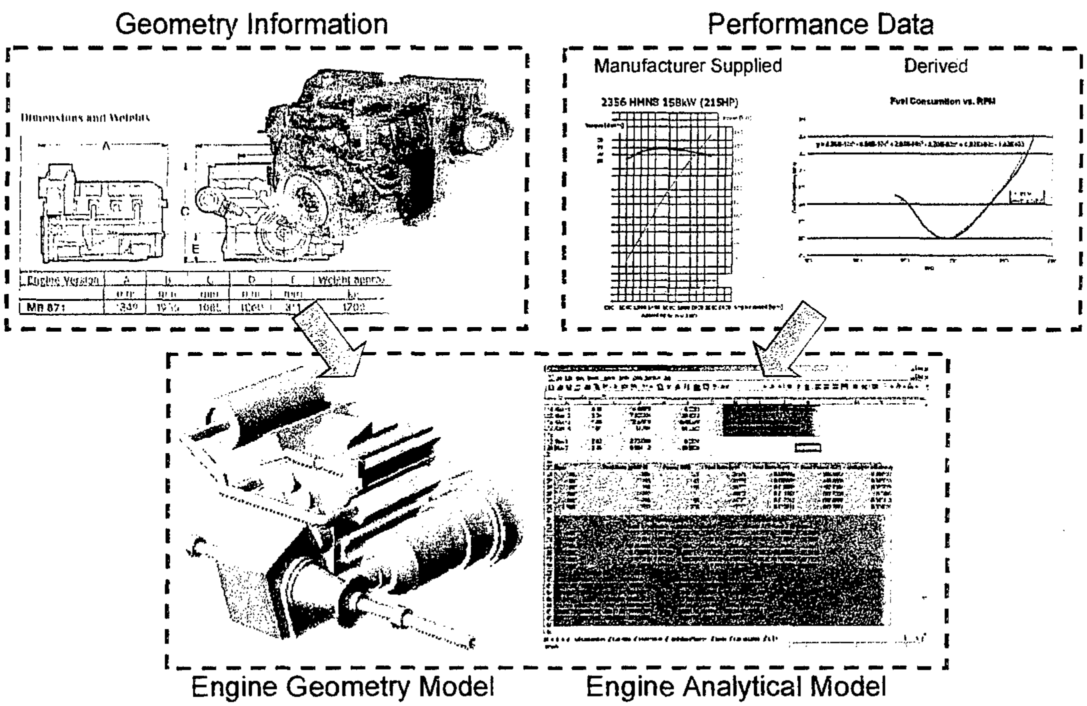


Figure 4. Engine Modeling for Ground Vehicle IR Signature Calculations

The engine modeling process begins with the collection of whatever data is available on the engine and drive train of the vehicle to be modeled. The first step is to create a three-dimensional model of the engine, generate a thermal mesh of the engine model, and incorporate the engine mesh into the vehicle thermal mesh. Analytical forms of the heat generation and heat transfer processes of the engine components must then be generated to create the software functions necessary to calculate the temperatures of the engine components. The code generation process begins with the collection of manufacturer-supplied automotive performance data or the derivation of these performance curves if

they are not available. This data is then assembled and curve fitted in spreadsheets to create the equations necessary to generate code. These equations are used to modify the PRISM source code, which is then re-compiled and executed to generate either the vehicle IR signature, or the necessary heat curves to drive MuSES.

Once the relatively difficult task of engine modeling is complete, the mesh is attributed in either PRISM or MuSES, simulation scenario parameters are set, and a thermal signature calculation is performed. It is important to note that the IR signature of a vehicle is a strong function of its environmental and operational history, therefore the temperatures of vehicle components must typically be calculated over an entire day in order to obtain the signature of a vehicle at a single time during that day. Figure 5 shows six example IR signatures of a BM-21. The signatures down the left-hand side of the figure represent a solar loaded daytime (13:00) signature in the three operational states of cold, idling, and exercised. The three signatures down the right side of the figure represent the same operational states for a nighttime (02:00) signature. All six of these signatures were generated with three PRISM or MuSES runs with signature models being output for two discreet times during the simulation.

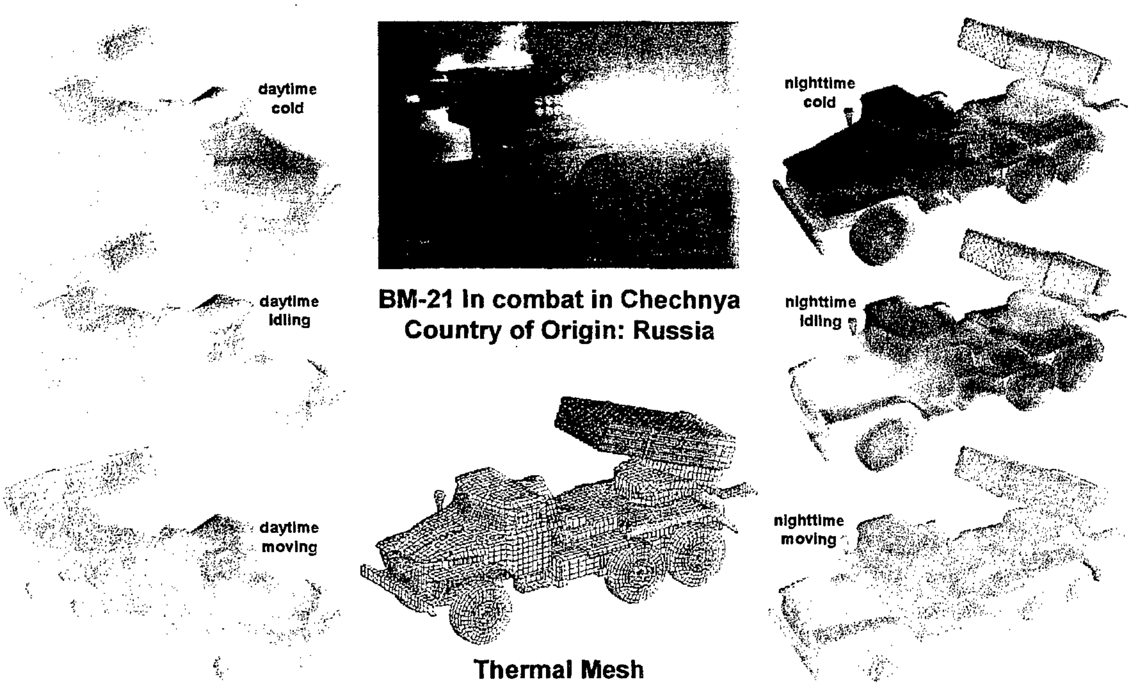


Figure 5. Example Ground Vehicle IR Signature Models

The different signatures shown in figure 5 demonstrate the power and flexibility of high-resolution predictive IR signature modeling. For example, the nighttime cold signature demonstrates passive radiative exchange between the vehicle model and the natural environment and between different sections of the vehicle itself. The flat surfaces that are exposed to the night sky appear colder and this is due to a net radiative heat loss to the sky from these components. Other sections of the vehicle that are shielded from the night sky, such as the area under the missile tubes, exchange thermal radiation with each other and stay relatively warm. The idling signature shows the effect of the engine

heating on the hood and other front sections of the vehicle, and the moving signature model demonstrates wheel heating and the effects of forced convection on the vehicle. Forced, ambient air convection causes the hood of the vehicle to cool off and the relative thermal contrast between other parts of the vehicle is reduced as the forced convection heat rate becomes relatively dominant and attempts to force the temperatures of many components to the ambient air temperature. The solar-loaded signature on the left hand side of figure 5 demonstrates self-shadowing as the thermal shadow of the rocket tubes can be seen on the bed of the vehicle and the roof of the cab.

Summary

Figure 6 shows the current high-resolution infrared signature model inventory available for generating signatures for the SIREEL website. All of these models were generated over approximately an 18 month period and more models are being generated at a very rapid rate thanks to the development of long lead time software and personnel infrastructure. The multipliers in the figure, for example "x6", mean that models of six major variants of that base vehicle have been produced.

The SIREEL website is currently available for IR combat identification training and it employs constantly evolving, state of the art techniques for generating infrared signatures of vehicles. Traditionally, ground vehicle signature models require many months to create. Due to the huge number of vehicles models that will be required for SIREEL, a highly efficient signature model assembly line has been set up and signature models are currently being created at a rate of two to three per month. Custom software tools have been created that allow for the rapid generation of real-time models from signature code outputs. The entire process results in a 3-D IR signature model that can be manipulated by a user in a web-browser over the Internet, and this initial technology implementation will be one of the basic tools used by SIREEL to train soldiers in infrared target recognition.

MBT's	**IFV's**	**APC's**	**MRL's**	**SPH's**	**Recon/Light Tanks**
T-55 (x6)	BMP-1	BTR-60	M-1991	M109	BRDM-2 (x4)
T-62 (x2)	BMP-2	M113	BM-21	2S1	PT-76
T-72		KIFV		2S3	
T-80 (x2)		MT-LB			
M1		BTR-50	**Anti Tank**	**AAA**	**SAM's**
M60			AT-3	ZSU-23/4	SA-9
K1			AT-5	ZU-23	SA-13
Chieftain			M-901	MT-LB/ZU-23	
				Gaz-66/ZU-23	

Figure 6. Q4 FY02 SIREEL IR signature model inventory

V & V of MuSES 6.0 Status Report

Teresa Gonda
TARDEC
Warren, MI 48397-5000

Al Curran, Alice Gerhardt, David Less, Pete Rynes, John S. Curlee
ThermoAnalytics 94X Airport Road
Calumet, MI 49913-0066

Bob Baratono
Keweenaw Research Center
1400 Townsend Drive, Houghton MI 49931

ABSTRACT

MuSES 6.0 was released late in 2001. TARDEC has undertaken a formal verification and validation procedure as described in the Army Modeling and Simulation Office Pamphlet 5-11 and as described in the Defense Modeling and Simulation Offices Guidelines on the subject. The full report is expected to be finished in the Fall of 2002. This paper will give examples of the latest results of the full environment Cubi tests and additional verification and validation tests. In addition, the paper will discuss some of the lessons learned on the V&V of engineering level models.

INTRODUCTION

The idea that everyone wants V&V is hard to dispute. What's more important however is that the DoD demands it. The Defense Modeling and Simulation Office (DMSO) released a comprehensive guide on the subject in 1996[1] and is currently updating it. In addition, they have established a Verification, Validation, and Accreditation (VV&A) technical working group to continue to mature the process, educate the community, and establish clear common nomenclature.

The Army Modeling and Simulation Office has followed suit and published guidelines under pamphlet PAM-5-11.[2] The official definitions of verification, validation, and accreditation in PAM-511 are as follows:

- "Verification is the process that determines the M&S functions as it was originally conceived, specified, and designed.
- Validation is the process that addresses the credibility of the M&S in its depiction of the modeled world.
- Accreditation is the M&S application sponsor's determination of the suitability and acceptability of the M&S to the application"

What VV&A can do is enhance a simulation's credibility and reduce the risk of its use in a particular application. What it cannot do is "guarantee that the modeling and simulation results will be correct, guarantee that the results will be correctly analyzed and interpreted, guarantee that the right model was chosen to solve the problem."[3]

UPHILL BATTLE

Even though program managers and engineers want validated models, very few organizations understand the process or want to go through the process. The reason is twofold. Firstly, it seems like an immense amount of work and

secondly, there is "Miller's Law". Certainly, the impact of picking up the thick Recommended Practices Guide on VV&A is to create a sense of being overwhelmed, but truthfully, the bulk of V&V it is just documenting good software development practices that should already be going on along the way. These procedures include creating requirements specifications, showing how requirements are tracked, documenting the sanity checks (verifications) and assumptions being made along the way, tracking configuration management procedures, and then finally the validation tests that any responsible proponent would perform on any model before they used it to their satisfaction. Once all of this is documented along with the strengths and the weaknesses of the model, then others can use this V&V report to judge whether the M&S is suitable for a particular problem. This "judgement" is the accreditation. Just because an M&S has weaknesses, does not make it invalid. If the weaknesses are understood and the M&S can still answer a certain question, then the M&S can be deemed appropriate and accredited for that purpose. Figure 1 lists some of the elements set aside for V&V in MuSES.

Assumptions made	Radiation variables	Signature inputs and outputs
Overall conduction	Environmental variables	Post processor features.
Overall radiation	All three modes of heat transfer	Geometry functions
Convection variables	Analysis of inputs and displays	Graphical features
Overall Convection	Editor functions	

Figure 2 Elements targeted for verification in MuSES

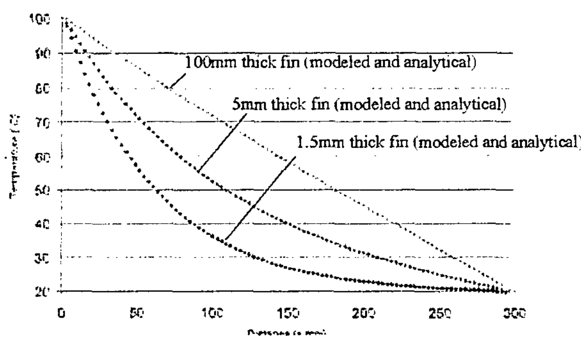


Figure 1 Results of conduction and convection validations

MILLER'S LAW

MuSES 6.0 has undergone the steps of individual V&V for the above elements. Comparisons to textbook analyses have shown agreement and the model behaves as expected. The final stage we are undertaking is face validation or output validation, which in this case is comparison to field trials using thermocouples and infrared imagers. This is the part of the process that most of those outside the V&V community think of as "validation". In the recommended procedures guide this part plays a small part in the over all process, but it is what most engineers and program managers wish to see. Unfortunately, this is also where Miller's Law comes into play.

Miller's Law states: Model *comparison to measured data is unfair*[4]

A portion of the sentiment behind this "law" is echoed in the DMSO guide. The very first principle listed in the guide is "There is no such thing as an absolutely valid model[1]". This principle addresses the idea that some might hold the mistaken belief that once a model was "validated" it was likely to give absolute answers that were "perfect". While there is a distinct wish to see plots of exact matches of modeled and measured data, under a great many circumstances that is unrealistic and not because a model is "bad". In fact, in certain cases, the modeled answer can actually be "more valid" than the measured due to uncertainties in measurements. This is the idea put forth in the adage, "no one believes a model except the one who ran it and everyone believes a measurement except the one who took it".

The reasoning behind Miller's Law is the following:

- Measurements are not always exact (due to inherent instrumentation uncertainties)
- Modelers are not always exact
- Modeling theories are often based on Measurements (see first bullet)
- Nature is not always well understood

Table 1 shows the reason as to why measurements are not always exact--in particular with respect to measuring "nature" or weather and environmental effects.

	Modeled – Measured Temperature Difference				
	Average	**Avg Abs**	**St Dev**	**Min**	**Max**
Baseline	-0.53	0.98	1.36	-10.6	6.64
Air T + 0.7 °C	0.061	0.86	1.31	-10.1	6.33
Solar Radiance + 5%	-0.26	0.91	1.29	-9.28	8.81
Solar Absorptivity + 0.05	-0.27	0.93	1.34	-9.94	8.56
Emissivity – 0.05	-0.39	0.93	1.36	-10.5	7.39
Cp 461 -> 420	-0.53	0.97	1.33	-10.4	6.72
Wind Speed – 5%	-0.49	0.99	1.35	-10.5	6.98
Air T + 1.0 °C	0.22	0.99	1.36	-9.84	7.02
Sky Radiance + 5%	0.16	1.00	1.42	-10.2	7.97
Thermal Conduct 52 ->73	-0.58	1.01	1.39	-10.8	6.32
Rotate Back 5 Deg	-0.61	1.02	1.36	-10.6	5.90
Insulated Back	-0.57	1.01	1.33	-10.3	7.82
One Layer with Air Inside	-0.62	1.05	1.50	-11.4	5.09

Table 1: Weather File Sensitivity Analysis, Output Validation

Miller's Law addresses the fact that in nature, there are many variables that fluctuate in nearly random fashion (with wind being the most dramatic) and the instruments that are used to measure nature and provide input data to models during validation exercises are not sensitive enough to capture these fluctuations. When this input data to the model is not accurate, it is unrealistic (unfair) to expect that models based on this input data will exactly predict what was then observed in the field.

Sometimes accurate data is theoretically possible, but very difficult to obtain in practice and determining how accurate data must be may be difficult to define. MuSES relies on material properties such as the emissivity and absorptivity broken out by wavelengths. This data set can be quite large, which can be a problem, but the bigger problem is that the data is hard to come by unless one owns the instrumentation to measure these properties. The measurement devices are expensive and the measurement itself is tedious--therefore there does not exist a large database of this type of information readily available. But the real problem with material properties is much more complex than this. The modeler needs to answer the question "what surface condition do I wish to model?" A lab measurement of a coating may be what is available for input into the model, but if one wishes to compare the results of the model to a measurement of a vehicle that has been in use for two years, the comparison could be very poor. Scratched and weather-beaten CARC paint for instance, has different properties than a coupon that is freshly painted and measured in a lab.

Miller's Law shines the light on the fact that face validation, while desirable, is a very difficult task to do well. The truth is that face validation has value and even if graphs of measured versus modeled do not match point for point, the conclusion is not automatically that there is a problem with the model. All modelers need to understand the assumptions made within the model and assumptions made about measurements used as input to the model. That said, we have proceeded forward with output validation (face validation) in the final stages of the V&V of MuSES. We will discuss the results, the examples of Miller's Law encountered, and the additional experiment designed to reduce its impact

OUTPUT VALIDATION

Since absolute comparison is not necessarily what is expected in output validation, how then can one measure a sense of "goodness?" When comparing measured and modeled data for MuSES, agreement is expected to fall within certain error bars calculated based on the above uncertainties. In addition, in these trials, we look at and compare temperature profiles over time. The results are time-stepped. Therefore, it can be said that it is also appropriate to look for agreement within plus or minus a given time step as well. Finally, there is "trend agreement." One can also visually tell if the predicted trend matches that of the measured trend. Under most circumstances, average agreement within plus or minus 2 degrees with good trend agreement is very good. Larger isolated deviations in and of themselves may not be a problem and are looked at on a case-by-case basis. Once modeling error, instrumentation error, and input data error is eliminated in these cases, algorithms and assumptions are analyzed. This is discussed in more detail later in this section.

Cubi

The geometry used in the output validation exercises of the MuSES/RadTherm code is called Cubi. Technion, in Israel first introduced this geometry in 1995 in order to validate PRISM. It was later adopted by the UK MoD. Liking the design, we chose the same geometry and dimension values in order to exchange data with our international partners. Cubi was purposefully designed such that it is fairly simple to model and troubleshoot, yet still presents some challenges to a model in a validation experiment. Firstly, past model validation experiments have shown that there is usually very good agreement in thicker armor. Therefore, in order to tax the model, thinner plates were used (which traditionally have been more difficult to match in predictions). Being box-like, terrain and sky effects are individually taxed/isolated via the different faces of the box. Then finally, the notch or step in the geometry allows faces to interact with each other (self-shadowing/self radiating).

Cubi is one meter tall by one half meter wide and one meter deep. It is constructed of 4.0-mm steel and was painted with a green CARC paint. Paint coupons were fabricated at the same time and sent out for material property measurements. Thermocouples were applied to the inside surface of the metal, and one-inch thick polystyrene foam board was pushed into place inside Cubi. The most recent thermocouple locations can be seen in the figure below.

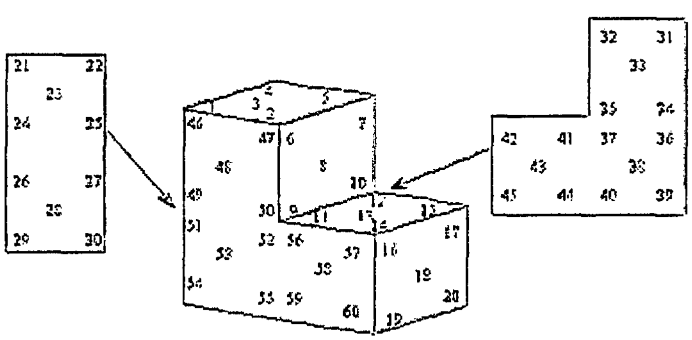


Figure 3 Cubi geometry and thermocouple map

Previous to this map, an initial test was performed in the spring in a small field close to the buildings at TACOM. There was a different thermocouple placement at that time. The results are in the figure below.

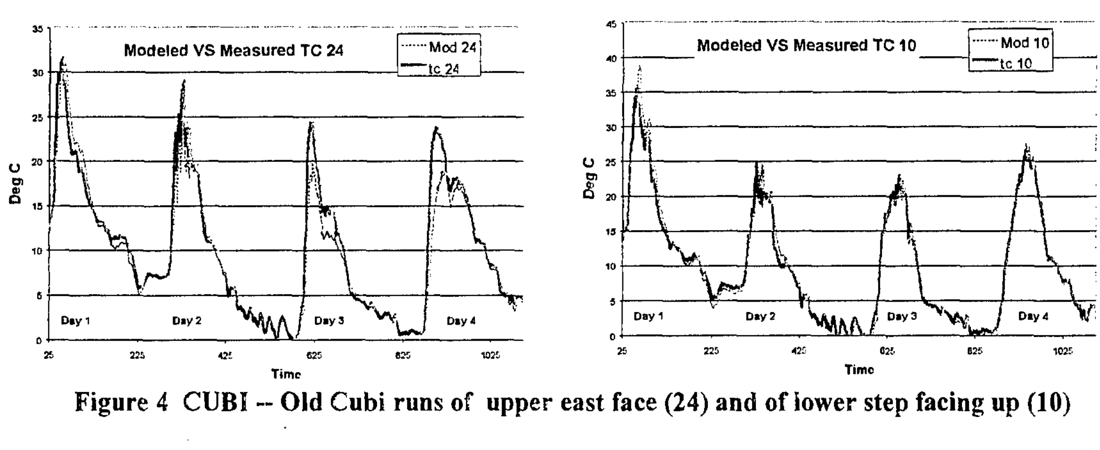


Figure 4 CUBI -- Old Cubi runs of upper east face (24) and of lower step facing up (10)

In this older trial thermocouple (tc) 24 is in the location of tc8 in figure 3 and tc10 is in the location of tc13. There are four days of consecutive comparisons. There is considerable agreement throughout the period with some problem areas during solar peak hours. A more rigorous second test was then designed to rule out testing errors and a new location was chosen that had less environmental clutter or interference from buildings and trailers that were not part of the experiment.

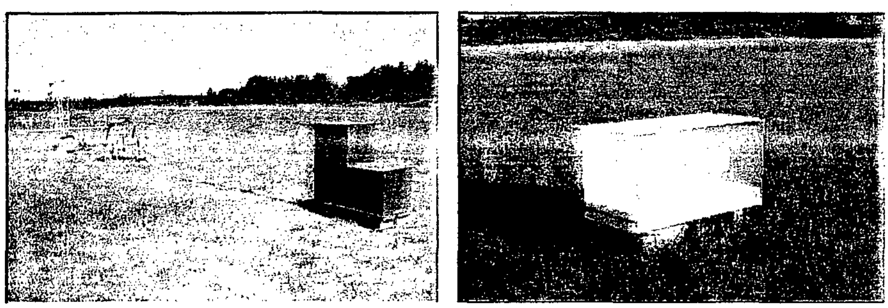


Figure 5 Cubi in the field and Plexiglas box for wind elimination experiment

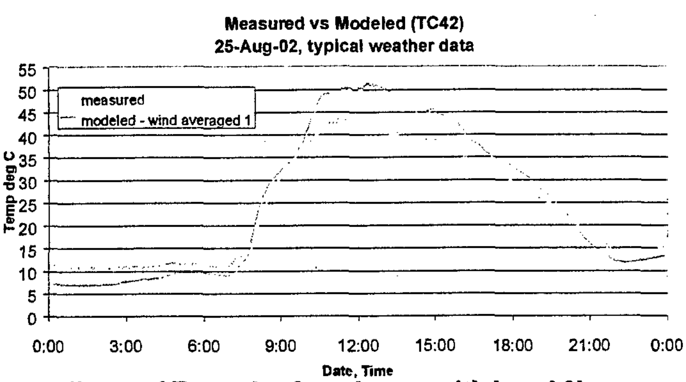


Figure 6 Example of results run with h and 2h.

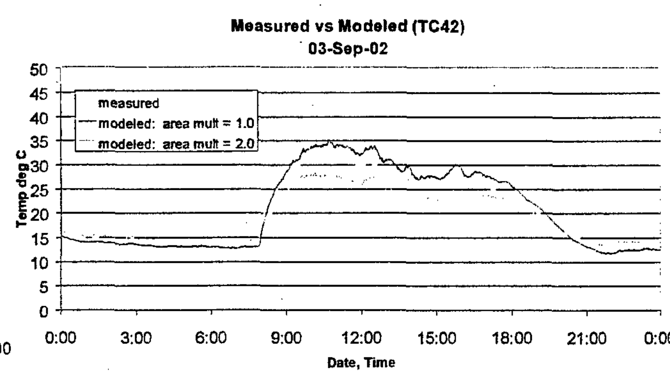


Figure 7 Example Cubi comparison in second trial

Problems with Wind

Ironically, in this experiment (which was carried out in an open field and therefore much more prone to wind) the comparison was worse than the less rigorous previous test. MuSES continually over predicted during the daytime solar loading period and under predicted at night. Based on ongoing discussions, experimenters began to suspect that wind convection was the problem. The convective approach used in MuSES had passed the V&V process up to this point, because it uses accepted textbook theory on the value of the convective coefficient h in wind and the associated Reynolds numbers. Those with experience in this area understand there are problems with this accepted theory. Miller's Law is the source behind the problem, because natural turbulent flows is one of physics unsolved mysteries at this juncture and theory is based on measured data[5]. Any comparison of measured to modeled data will be "unfair" where turbulent wind is a factor, because there is an inherent approximation or error built into the theory. This is because measuring wind out in nature is difficult and therefore theories are developed based on the more predictable continuous flow wind tunnels that can be measured more effectively. But we also know that due to the turbulence created by high speed fluctuations of natural wind, the real value of h (and hence, the effective cooling) is higher than the smooth wind tunnel flows. There is however, no exact theory for calculating the exact speed at which wind moves from laminar to turbulent nor the exact Reynolds number to use to calculate h once it does. It can be as much as 2-3 times more than the accepted textbook value. Using this example, it is perfectly "valid" to understand that using accepted theories, predicted values will result in temperatures that are higher than seen in the field. However, outside of implementing an exhaustive computational fluid dynamics capability within MuSES, we wish to

allow the user to get a value more in line with what would be expected in the field. We have determined that we need a better approach. Firstly, to validate that this was indeed the issue, trial and error was used to determine if there could be found a "right" value of h to fit the data. Figure 7 shows the result of standard h and doubling the value of h. The result is that the model now has reversed the trend--under predicting during the day and over predicting during the night. Clearly, in this case, 2h is over compensating. The actual value in this example was a little over 1.5. The next step in validating that this was the problem with the matching was to design an experiment to eliminate the effects of wind and see if this removed the problem. A Plexiglas container was built (figure 5) and placed over a flat plate with a solar sensor and an aspirated air sensor placed inside. Some example results are shown below.

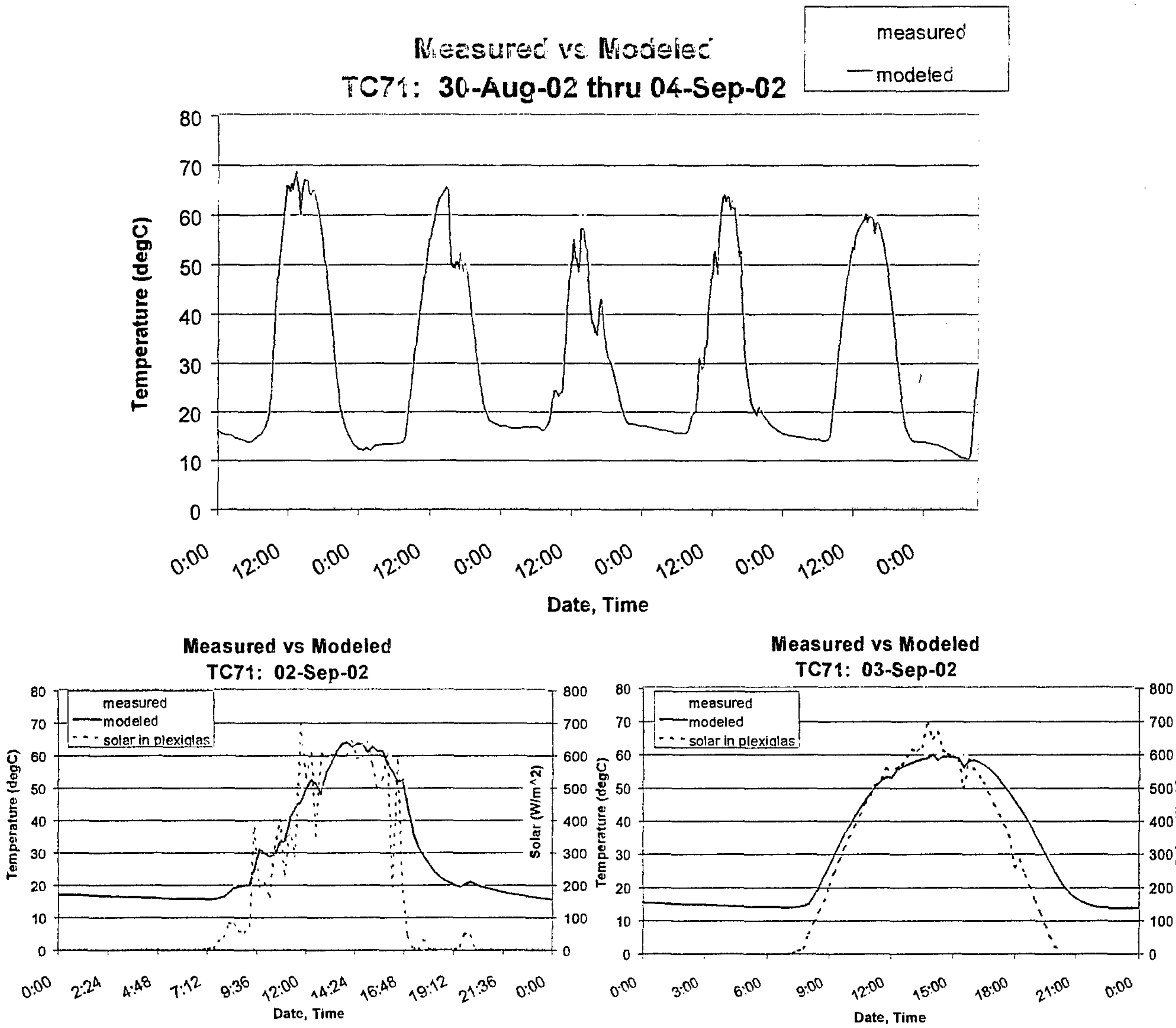


Figure 8 Output from Summer 2002 Covered Plate Trial (a-upper, b-lower left, c-lower right)

This sample is the worst of the data and yet clearly the agreement is quite good--supporting the earlier conclusion. Figure 8a shows five straight days of measurements, while 8b and 8c are close-ups of particular days. There is a small error inherent in the experiment at between 1500h and 1600h, where the solar sensor's view is blocked by corner in the Plexiglas container, effecting the modeled results. Notice the absolute agreement during non-solar hours. In addition, there is some departure during times of fluctuating cloudiness (where cloud solar scatter affects the validation experiments). To aid in analysis, the measured solar data is also plotted to track cloud activity and the blocked solar sensor. These runs were executed with the more simple weather file carried over from PRISM that uses a cloud cover number that is an approximation by an observer.

Additional runs will be made using a more rigorous weather scenario, but for now, the conclusion is clear. The theory might be acceptable, but in order to provide the user with a more robust prediction in nature, an alternative approach will have to be developed for inclusion in MuSES 7.0 (See Lessons Learned Section)

Apparent Temperature

In addition to physical temperatures, validation tests are scheduled for apparent temperatures as well. Figure 9 shows some data from a previous experiment where MuSES predictions of apparent temperature were compared against data taken from imagery. Taking into account temporal trends (how close the data points are to the line, not just in the y-axis) that are much easier to see in the plot on the right, the data falls within the aforementioned 2 degrees C except for one point. The standard deviation also falls within this range and the root mean square (RMS) value is 1. The one-off larger deviation was accounted for in the experiment. More exacting apparent temperature experiments will be conducted using Cubi.

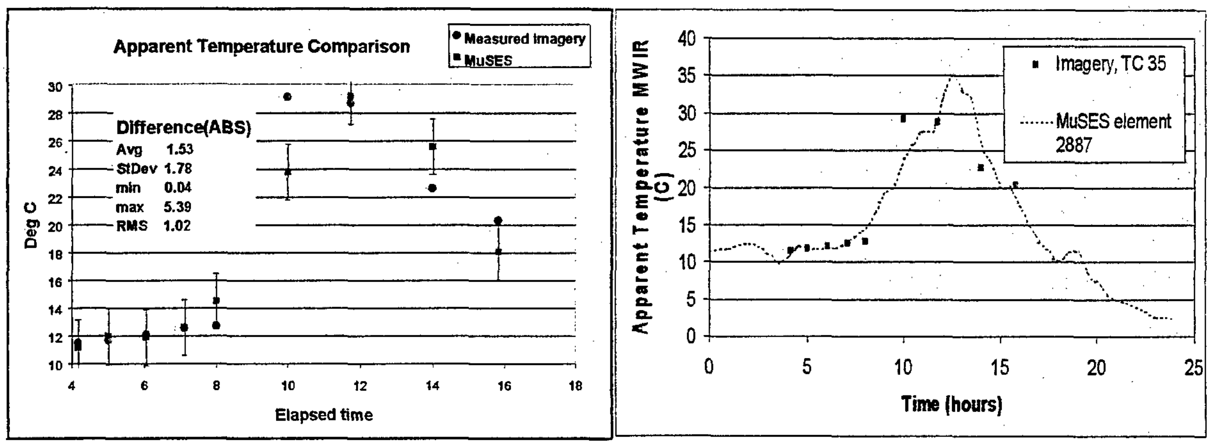


Figure 9 Apparent temperature comparison

LESSONS LEARNED

During the initial test and the second test, many lessons were learned by both persons new to MuSES and by experienced professionals. Part of our V&V process included having new users exposed to the output validation process to investigate problem areas not just with MuSES, but with the modeling experience and in the process of comparing measured and modeled results. Some of the lessons learned were obvious some were not.

New Wind Model

One outcome of the second test is that MuSES 7.0 will have a new wind model. In addition to accepted theory that is known to have its limits, the user will be given a choice of additional approaches. A new model will be developed using the data from the second Cubi/plates experiment. This high quality data, while limited in geometry shape, provides a great deal of new information on flat plates facing in different directions in multiple wind speed situations. This new model will be derived using time of day, plate angle, wind direction and speed and will be compared with the accepted theoretical formula and work of previous researchers.

Material Properties

The materials database in MuSES should be understood to contain <u>representative</u> values for listed items. Depending on the question being asked of the model, this choice might be appropriate. In a validation experiment or if the answer needs

to be exact. one has to obtain measurements of the properties. For a validation experiment. one has to know the exact emissivity and absorptivity of the surface of the test object and the measurements need to be made on a coupon that is in the same condition as the validation object measured in the field. Sensitivity analyses show that this area is critical. Lesson: If absolute answers are important it is worth the time or expense to get coupons made and measured. In addition, if the trials potentially will last some time out doors, have coupons measured before the test begins and then keep a coupon with the test set up, so measurements can be made at the end of the trials to check for changes.

Diligent Monitoring of Meteorological Data

It is crucial that the test and measurement equipment be set up properly. The weather station used in the early Cubi trials measured both the shadow band diffuse solar radiance and the total solar radiance. Upon examination, it was determined that the equipment used to measure diffuse solar was allowed to get out of alignment, causing contamination of the measurements by direct sunlight in the morning hours. Lesson: It is critical to pay close attention to this area. In this second phase, a person was put solely in charge of developing, calibrating, and maintaining the met station as well as data reduction. Bad meteorological data can invalidate days of data.

Thermocouple Data

Not only is it crucial that test and measurement equipment be set up properly, but it is also necessary to build in redundancy to maintain good data and mitigate risk against malfunctioning equipment whenever possible. Towards this end, five thermocouples were applied to each face of Cubi. This was planned to help capture gradients along the faces, but more importantly to provide redundancy. This proved valuable, since having more than one thermocouple on each face made it very easy to discover that thermocouple 14 had malfunctioned during the first Cubi trials.

Invalid comparisons of modeled and measured.

A common problem, particularly with inexperienced modelers, is comparing modeled and measured data, where not all the relevant information is put into the model. Often in an attempt to create short cuts, certain assumptions are made to save time. For example, it was determined that in the first Cubi test the plywood base (which extended out past the bottom of Cubi by approximately an inch) actually interacted enough with the model to make a difference. The test location itself was also questions--there was an uncertainty as to whether buildings and trailers in the test vicinity had much of an impact on the test results. This exact problem occurred in Israel's early Cubi tests as well. Comparisons to measured and modeled data were problematic, but the test setup was on a rooftop with building and instrumentation surrounding the test object--these interactions were not included in the model. Lesson: Pay close attention to reduce interactions with items that are not part of the experiment or take the time to model them accurately. Assumptions of non-interaction are often incorrect.

Cloud Scattering Spikes

When taking validation weather data, in is important to be cautious of partly cloudy days and the solar measurement readings. Per Dr. Ann Webb at UMIST, "The phenomenon is well known amongst those who routinely measure solar radiation, and is apparent when measuring both total (all wavelength) radiation and discrete spectral. Radiation travelling thorough the atmosphere is scattered whenever a photon happens to meet a particle: some photons (direct beam radiation) reach the earth uninterrupted. The scattering falls into a number of regimes. Clean air particles provide Rayleigh scattering that goes half in the forward (earth) direction and half backwards (to space). Other constituents of the atmosphere (pollution, cloud droplets) are much bigger and their ratio to the size of the wavelengths of solar radiation put them in the Mie scattering regime. Mie scattering patterns change with the ratio of wavelength to particle size, but generally produce a stronger forward scattering compared to the Rayleigh case. As a cloud comes close to the path of the direct solar beam and circumsolar radiation, this strong region of radiation is subjected to forward scattering ... and an increase or spike in the data record is recorded that can be substantially more than the clear sky value. Usually the cloud then passes in front of the sun and the direct radiation is blocked, causing a significant drop in signal. As the cloud clears the solar disk a further high radiation value is briefly recorded. The net effect of all this is that cloud does decrease radiation reaching the earth (as one would expect), but for transient moments it can increase values above those of clear skies, especially on summer days with scattered fair weather cumulus ("cotton wool clouds")."[6]

As these examples show it is the user's responsibility to determine the appropriate level of accuracy required in any model and simulation and when the resolution is "good enough". This is why we are running sensitivity analyses on the code. In this way, modelers can better determine how much attention needs to be paid to individual parameters in order to achieve predictions within a certain tolerance.

SUMMARY

MuSES/RadTherm is undergoing a great deal of scrutiny by both industry and the government. MuSES 6.0 is proving out to be very robust. While not technically a bug in V&V terms, the wind model needs improvement and there is already a change set for version 7.0 to be released in the late Fall of 2002. The "comparison with measured data" tests, on which many people hang their hats, have proven successful to this point. Additional apparent temperature trials will be conducted using Cubi and V&V will be an ongoing process with each new release--scheduled every six months.

As is the case with other sophisticated engineering level software tools, users need to receive training in the software and need to have at least an intermediate understanding of heat transfer in order to best utilize the code. And as is the case with all models, the user must know what question he or she wants the model to answer and if the model is capable of answering it. At this stage, (i.e. within the bounds of what has been tested) MuSES/RadTherm 6.0 performs within acceptable tolerances and gives expected results.

Unless a modeler is doing a validation or a model-test-model approach, comparison to measured data will most likely not be possible. It is our hope that by these lessons learned and the data contained within, that modelers may better be able to determine the most effective method of modeling for a particular situation.

ENDNOTES AND ADDITIONAL REFERENCES

[1] DMSO "Verification, Validation and Accreditation (VV&A) Recommended Practices Guide, 1996, https://www.dmso.mil/public/transition/vva/; current guidance: https://www.dmso.mil/public/transition/vva/policiesguidance; plus the updated DoD instructions: https://www.dmso.mil/public/library/projects/vva/products/DoDI%205000.61%20-dbl%20spa%20draft_7-10-2002a.doc

[2] http://www.amso.army.mil/; http://www.amso.army.mil/library/index.htm

[3] DMSO "Verification, Validation and Accreditation (VV&A) Recommended Practices Guide, 1996

[4] Brian Miller, research engineer, CECOM, Night Vision Electronics and Sensors Directorate (NVESD).

[5] Mills, A. F., Basic Heat and Mass Transfer 2/E, 1999 Prentice Hall, Inc., p382

[6] Per email, Dr Ann Webb, a solar researcher at UMIST in England.

Additional References:

http://www.informs-cs.org/

http://www.acm.org/pubs/tomacs/

http://www.sisostds.org/

A CONCEPT FOR ASSISTING REMOTE IR DAMAGE ASSESSMENT BY COUPLING NUMERICAL FIRE SUPPRESSION MODELS WITH THERMAL SIGNATURE MODELS

Neal E. Blackwell, Ph.D
Army Power Division
U. S. Army CECOM RD&E Center (CERDEC)
Fort Belvoir, VA 22060-5816

Judy Cooper
AEA Technology

Steven J. McCormick
U. S. Army TARCOM RD&E Center (TARDEC)

Jeffery V. Mosley
OptiMetrics, Inc.

ABSTRACT

Program Managers (PMs) under the Program Executive Office for Ground Combat Systems (PEO-GCS), funded the U.S. Army CECOM RD&E Center (CERDEC) to develop and apply a 3D numerical model to predict the performance of automatic fire suppression systems in armored combat vehicles with Halon and Halon alternative fire suppressants. Running the model without suppression allows predictions of fuel spray fire propagation in defeated vehicles and the resulting internal temperatures. The model also predicts the extinction of the fire due to oxygen availability. Hence, transient heat flux to interior surfaces can be predicted until fire extinction and can subsequently used as boundary conditions for a thermal signature code such as PRIZM or MUSES.

A numerical fireball model specially formulated for fire suppression is incorporated into a 3D, transient, hybrid, finite-volume, finite element code with transient tracking of JP8 and fire suppressant particles. The numerical fireball model consists of a flame front submodel driven by finite-rate kinetics that describe the reactions between sprayed JP8 fuel particles, oxygen and suppressant particles. The purpose of the numerical fireball model is to predict fires resulting from fuel tank or pressurized fuel line penetrations.

This information is reported in order to propose the use of the PEO-GCS Fire Suppression Model as input to thermal signature prediction models. The proposed result of this concept would be to gain the ability to predict thermal signatures of defeated enemy armored vehicles for comparison with thermal images from remote IR sensors for damage assessment purposes. This concept would be most valuable when major exterior structural damage (turret separation, etc.) is not present as evidence for damage assessment purposes.

INTRODUCTION

Damage assessment of armored vehicles is often challenging when there is a lack of exterior clues such as turret separation or large soot coverage. Hence, this paper proposes a concept to assist those involved in damage assessment of armored vehicles where few exterior clues exist. The concept uses time lagged evidence of a large internal heat release due to catastrophic fires in armored vehicles. The resulting time-temperature history of vehicle armor is lagging in time due to the heat capacity of armor. For damage assessment purposes, the advantage of high heat capacity in metallic armor is that it can store large amounts of heat which are then released slowly to the atmosphere and hence can be sensed by infrared sensors for relatively long periods of time

METHODS

Time Lagged Armor Heat Release

Thermal analyses involving unclassified work on armored vehicles often use the assumption that the thermal properties of special recipe armor can be approximated by the thermal properties of lead. Additionally, the lumped heat capacity technique can be used in transient thermal predictions where the surface convection resistance is very large compared with the internal conduction resistance. A criterion to test this assumption is the parameter from Chapman (1974),

$$hm/(\rho Ak) < 0.1 \tag{1}$$

where: h = convective heat transfer coefficient,
m = vehicle mass,
ρ = armor density,
A = exterior surface area of armor, and
k = thermal conductivity of armor.

The value of this criterion is approximately 0.005 and 0.004 for lead-like and aluminum-like armor recipes. Hence, this assumption was used to derive the following first order, linear, ordinary differential equation which expresses, in general form, the armor temperature resulting from an interior fire:

$$dT/dt + ZT = K \tag{2}$$

where: T = armor temperature,
t = time,
$Z1 = (1/c_p m)*(h_i A_i + h_\infty A_\infty)$,
$K = (h_i A_i /(c_p m))T_i + (h_\infty A_\infty /(c_p m))T_\infty$,
Ti = combustion gas temperature adjacent to interior walls
$T\infty$ = outside air temperature
c_p = heat capacity of armor,
A_i = interior surface area, and
A_∞ = outside surface area.

The solution of equation 2 was found using the method of separation, which yielded the following expression for the armor time-temperature history, during the fire:

$$T = (T_o - K/Z)e^{-Z1t} + K/Z \tag{3}$$

where: T_o = initial temperature of armor

Equation 4 below from Holman (1976), mathematically describes the armor cooling period after the fire is extinct,

$$T = (T_o - T_\infty)e^{-Z2t} + T_\infty \tag{4}$$

where: $Z2 = (h_\infty A_\infty)/(c_p m)$.

Resulting predictions are for the conditions of, 1) 300 seconds of heat from fires inside of 20 ton and 40 ton armored vehicles, in the Army Hot Dry (Desert) condition of 49C (120F), 2) a 10 mph wind, 3) desert sand background temperature of (60 C)140F, and 4) an initial armor temperature equal to the desert sand background temperature. This study reports predictions that are bounded by the assumption that the fire will heat the gases adjacent to the interior walls to temperatures between the lower and upper bounds of 400 C (750F) (673K) and 1000 C (1830 F) (1273 K) respectively, with a mid-bound of 700 C (1292 F) (973 K). The combustion gas and fire temperature bounds were chosen to agree with measurements conducted by the Army's Aberdeen Test Center and reported in Gritzo et al. (1999). These predictions also

assume that the fire intensity remains constant throughout the 300 second (5 minute) fire exposure period. Predictions for shorter fire exposure times can be extrapolated from Figures 1, 2 and 3.
This study used an example from Driggers et al. (1999) of a Second Generation Forward-Looking Infrared (FLIR) imager and a standard 2.3 x 2.3 m NATO target with a 1.25 C target-background temperature difference. The conditions were 1) clear. 2) the standard U.S. atmosphere and 3) an altitude of 0.5 km above sea level. Values for the 50% number of cycles across target N_{50} for identification, recognition and detection were 0.75, 3.0 and 6.0. respectively. For this example, Driggers et al. (1999) predicted that the probability of detection is greater than 0.9 for distances up to 5 km (3.1 miles). Much greater detection distances are expected for the relatively large target-background temperature differences considered in this study.

Combining the prediction by Driggers et al. (1999) with the predictions from Equations 3 and 4 above, a 20 ton target with a 1.25 C target-background temperature difference is detectable up to 1.0, 1.5 and 1.8 hrs, for interior combustion gases of 400, 700 and 1000 C, respectively, for a 300 second (5 min) fire (Figures 1, 2 and 3). A 40 ton target with a 1.25 C target-background temperature difference is detectable up to 1.4, 2.1 and 2.5 hrs, for interior combustion gases of 400, 700 and 1000C, respectively, for a 300 second (5 min) fire (Figures 1, 2 and 3).

Driggers et al. (1999) reported that typical, ground target, target-background temperature differences were in the range from 1.25 to 4 C. Therefore, thermal signatures resulting from internal fires can only be distinguished for target-background temperature differences greater than 4 C. For a 20 ton target, the target-background temperature difference is greater than 4 C for up to 0.9, 1.3 and 1.6 hrs. for interior combustion gases of 400, 700 and 1000C, respectively, for a 300 second (5 min) fire (Figures 1, 2 and 3). Likewise, for a 40 ton target, the target-background temperature difference is greater than 4 C for up to 1.0, 1.7 and 2.2 hrs, for interior combustion gases of 400, 700 and 1000C, respectively, for a 300 second (5 min) fire (Figures 1, 2 and 3). Hence, it seems plausible that the time periods over which defeated vehicles have a target-background temperature difference over 4 C are long enough to provide an interior fire signature for assessment and comparison with predicted signatures.

Figures 1, 2 and 3 also show that, in the event of such catastrophic fires as are considered in this study, light armored vehicles are predicted to have greater target-background temperature differences than more heavily armored vehicles. Damage assessment IR surveillance equipment can therefore detect these signatures from light armored vehicles with greater probability and at greater distances than from heavily armored vehicles. However, the window of opportunity to detect these thermal signatures from light armored vehicles is shorter than for heavily armored vehicles (Figures 1, 2 and 3), making timely damage assessment more crucial in these cases.

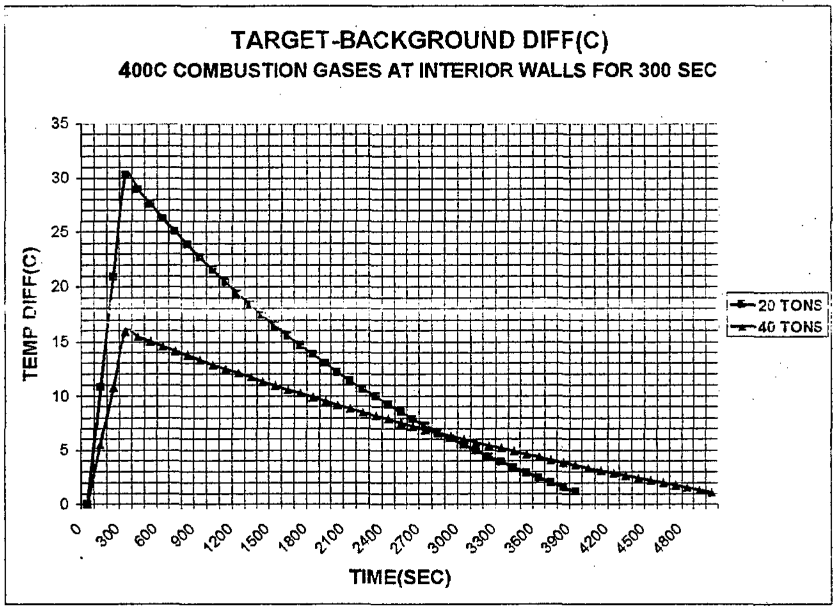


Figure 1. Target-Background Temperature Difference for lower bound of fire intensity exposing 400 C (750F)(673K) hot gases in the interior walls of a 20 ton and a 40 ton armored vehicle for a 300 second fire exposure time.

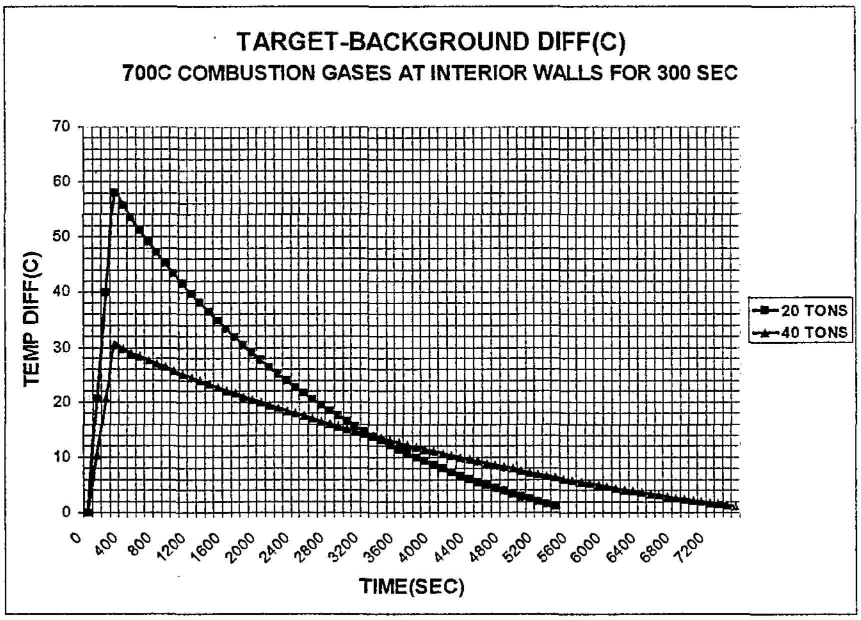


Figure 2. Target-Background Temperature Difference for lower bound of fire intensity exposing 700 C (1292 F)(973 K) hot gases in the interior walls of a 20 ton and a 40 ton armored vehicle for a 300 second fire exposure time.

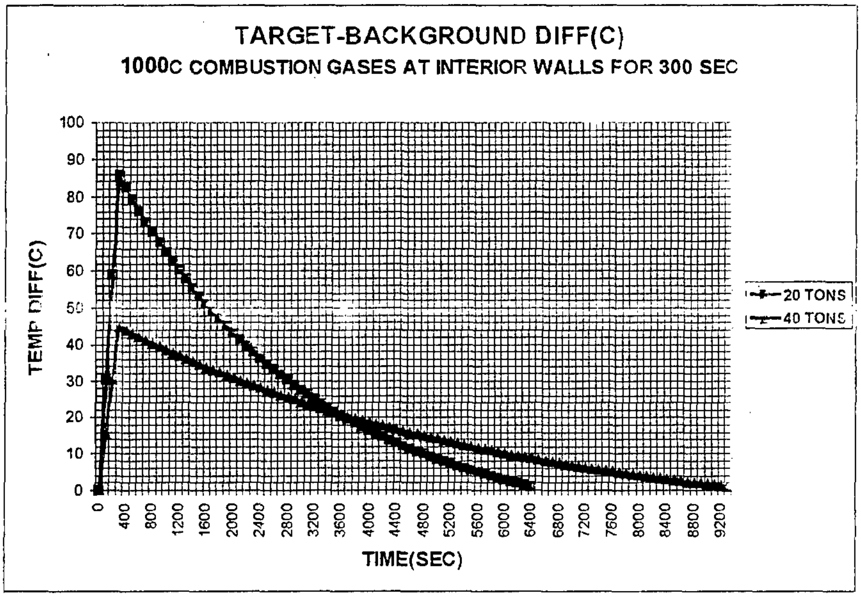


Figure 3. Target-Background Temperature Difference for lower bound of fire intensity exposing 1000 C (1830 F) (1273 K) hot gases in the interior walls of a 20 ton and a 40 ton armored vehicle for a 300 second fire exposure time.

Program Managers (PMs) under the Program Executive Office for Ground Combat Systems (PEO-GCS), funded the U.S. Army CECOM RD&E Center (CERDEC) to develop and apply a 3D numerical model to predict the performance of automatic fire suppression systems in armored combat vehicles with Halon and Halon alternative fire suppressants (Blackwell et al. 2000, Blackwell et al. 2001a, Blackwell et. al. 2001b, Blackwell et al. 2002a, Blackwell et al. 2002b, Skaggs et al. 2002). Running the model without suppression allows predictions of fire propagation in defeated vehicles and the resulting internal temperatures. The model also predicts the extinction of the fire due to oxygen availability. Hence, transient heat flux to interior surfaces can be predicted until fire extinction and subsequently used as boundary conditions for a thermal signature code.

This information is reported in order to propose the use of the PEO-GCS Fire Suppression Model as input to thermal signature prediction models. The proposed result of this concept would be predicted thermal signatures of defeated enemy armored vehicles for comparison with thermal images from remote thermal sensors for damage assessment. This concept would be most valuable when major structural damage (turret separation, etc.) is not present for damage assessment purposes.

Numerical Code Description

A numerical, conservative, hybrid finite volume-finite element formulation was used to solve the steady state equations for mass, momentum, energy and the transport equations for turbulent kinetic energy and the dissipation of turbulent kinetic energy. The commercial code is licensed by AEA Technology under the name of CFXTASCflow and has multi-block, body-fitted coordinate and local grid refinement capability.

The differencing scheme used for the advective terms was the mass-weighted scheme in conjunction with Physical Advection Correction which is formerly 2nd order accurate (Huget, 1985; Raithby, 1976; Lillington, 1981; Raw, 1985). The 2nd order, Central Differencing scheme was used for the diffusion terms. For continuity, pressure velocity correction was solved in a coupled manner and a modified Rhie Chow Interpolation was used to smooth out numerical oscillations in the solution and yield an error on the order of the 4th derivative.

Transient spray particles of fuel were modeled in the Lagrangian frame of reference and the gas species and reactions were modeled in the Eulerian frame of reference. The fire is computationally ignited with a small volumetric heat source of 20 to 30 kW near the ballistically created hole from which fuel spray is injected into the domain.

Most of the commercial code is parallel for shared memory machines and distributed machines. The relevant submodels for fire predictions that are not parallel are the submodels for, 1) transient particle tracking of fuel particles, and 2) fire. Future plans include the parallelization of these capabilities for PC clusters and for shared memory machines which are common in DoD High Performance Computing Centers.

Suppression/Combustion Submodel Description

The commercial, numerical code was customized to predict fire propagation resulting from a ballistic threat through a fuel tank or pressurized fuel line. Code enhancements relevant to the concept discussed herein include, 1) JP8 spray fires using the Eddy Dissipation Concept (EDC) submodel, with modifications by Rasmussen and Myken (1994), in conjunction with the flame front submodel with flame speed modifications derived from data by Linteris and Truett (1995) and 2) a two reaction mechanism for JP8 combustion by Westbrook and Dyer (1981)(Blackwell et al. 2000). The flame front submodel is driven by finite-rate kinetics that describes the reactions between sprayed JP8 fuel particles, oxygen and suppressant particles. Turbulence was modeled using the two equation k-e model (Launder and Spalding, 1972 and 1974). Radiation was modeled using the Gibbs model, which is applicable to opaque domains. A discrete ordinance radiation model was available but was judged to be too computationally expensive in comparison to the Gibbs model, which is acceptable for opaque domains, such as has been observed in recent ballistic tests by Skaggs (2003). Recent high speed video of ballistic tests by Skaggs (2003) compared with high speed video of spray generator initiated fires by ATC (1999), show the domain is much more opaque in the ballistic event. The author postulates that the increased opacity resulting from the ballistic event is due to 1) soot, 2) spall, 3) behind armor debris and 4) other ballistically airborne particles. Soot was modeled using the Magnussen soot model. Fuel spray boundary conditions for the PEO-GCS Fire Suppression Model are predicted using the Fire Prediction Model developed by Andy Pascal of Enthalpy, Inc. for the Air Force or by using spray generator data that was measured and reported in Reed et. al. (2002).

DISCUSSION

Program Managers (PMs) under the Program Executive Office for Ground Combat Systems (PEO-GCS), funded the U.S. Army CECOM RD&E Center (CERDEC) to develop and apply a 3D numerical model to predict the performance of automatic fire suppression systems in armored combat vehicles with Halon and Halon alternative fire suppressants. Running the model without suppression allows predictions of fuel spray fire propagation in defeated vehicles and the resulting internal temperatures. The model also predicts the extinction of the fire due to oxygen availability. Hence, transient temperatures of combustion gases adjacent to interior surfaces can be predicted until fire extinction and can subsequently used as boundary conditions for a thermal signature code such as PRIZM or MUSES. Figure 4 shows an example of output from the PEO-GCS Fire Suppression Model, where hot gas temperatures adjacent to crew faces are predicted versus time. Time-temperature histories at the interior walls could be stored and input as boundary conditions in a numerical thermal signature code like PRIZM or MUSES. The proposed result of this concept would be gain the ability to predict thermal signatures of defeated enemy armored vehicles for comparison with thermal images from remote IR sensors for damage assessment purposes. This concept would be most valuable when major exterior structural damage (turret separation, etc.) is not present as evidence for damage assessment purposes.

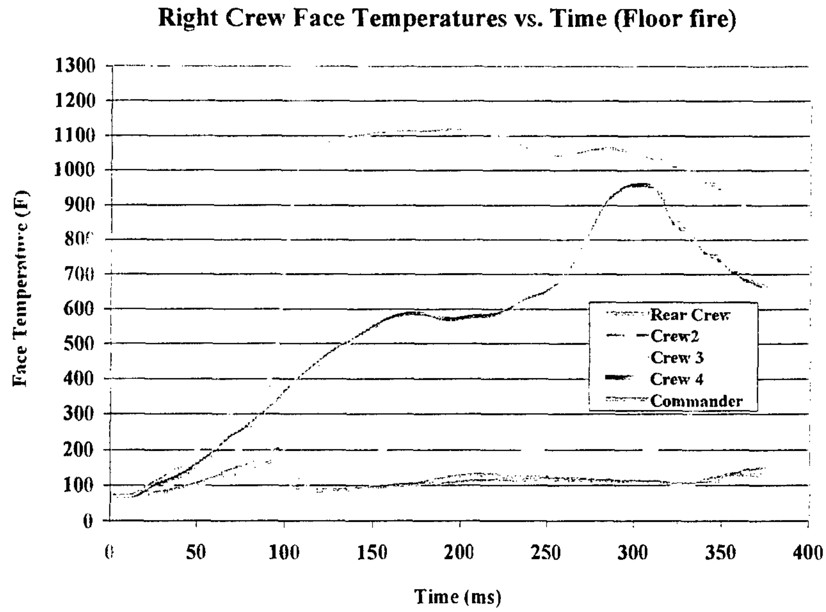


Figure 4. Predicted Crew Facial Temperatures versus time (Blackwell et. al., 2002b).

CONCLUSIONS

The existing PEO-GCS Fire Suppression Model could provide transient, thermal boundary conditions to numerical thermal signature models of enemy armored vehicles for the purpose of predicting the thermal signature associated with an internal catastrophic fire, hence, indicating a kill. This concept would be most valuable when major exterior structural damage (turret separation, etc.) is not present for visual damage assessment.

ACKNOWLEDGEMENTS

The author wishes to thank 1) Mr. Theodore Vician, formerly of PM Bradley, 2) Mr. Steve McCormick, head of the TACOM Halon Replacement Program, 3) Mr. Robert DeGroot, formerly head of Systems Integration for PEO GCS, 4) Mr. Richard Hutchinson, head of AAAV Survivability for the US Marine Corps and 5) Mr. John Karavias of PM Abrams, for championing a consortium to fund the initial development of this numerical fire suppression model for armored vehicles. The author also wishes to thank Major General John F. Michitsch, former PEO for GCS, for his interest and support of this effort in the context of the Army Halon Replacement Program. Gratitude is expressed to Mr. Al Schumaker, Mr. Gary Schultz, Mr. Terry Dean and Mr. Jeff Mosley of the PM BCT Office for supporting the recent STRYKER fire suppression modeling work. Finally, appreciation is expressed for the technical expertise of Dr. Kevin Knill and Dr. Ross Haywod, formerly of AEA Technology and Mr. William Elliott of Kidde Dual Spectrum.

REFERENCES

ATC. 1999. Fireball Generator-Poverty Island. U.S. Army Aberdeen Test Center. Aberdeen, Maryland.

Blackwell, N.E. and Cooper, J., Haywood, R., and Knill., K. 2000. Virtual Fire Suppression: Consortium Results – BRADLEY, AAAV, ABRAMS. PEO GCSS, PM Bradley, PM AAAV, PM Abrams. U.S Army Communication-Electronics Command Research, Development and Engineering Center, Fort Belvoir VA..

Blackwell, N.E., Vician. T.E., and McCormick, S.J. 2001a. Virtual Fire Suppression in Combat Vehicles with Halon Alternative Fire Suppressants. Unclassified Proceeding of the 12th Annual U.S. Army Ground Vehicle Survivability Symposium. U.S. Army Tank Automotive Command, Warren, MI.

Blackwell, N.E., Cooper, J., Connaghan, M.D. 2001b.CFD Analysis of Fire Growth and Suppression in the AAAV. U.S Army Communication-Electronics Command Research, Development and Engineering Center, Fort Belvoir, VA, 2002.

Blackwell, N.E., Cooper, J., Connaghan, M.D., Mosley, J., and McCormick, S.J. 2002a. Virtual Fire Suppression for Future Combat Vehicles. Unclassified Proceeding of the 13th Annual U.S. Army Ground Vehicle Survivability Symposium. U.S. Army Tank Automotive Command, Warren, MI. – in submission, 2002.

Blackwell, N.E., Cooper, J., Connaghan, M.D., McCormick, S.J. and Mosley, J. 2002b. CFD Analysis of Fire Growth and Suppression in the STRYER ICV. U.S Army Communication-Electronics Command Research, Development and Engineering Center, Fort Belvoir, VA, 2002.

Chapman, 1974. Heat Transfer. 3rd Edition, Macmillan Publishing Co., Inc., New York, New York.

Driggers R.G., P. Cox and T. Edwards. 1999. Introduction to Infrared and Electro-Optical Systems. Artech House, Inc., p 317-319.

Gritzo, L.A., A.R. Lopez, W. Gill and J.M. Williams. 1999. Numerical Modeling of Fires in Ground Vehicle Crew Compartments: Calculation Results and Comparison with Experimental Data. Sandia National Laboratories, Albuquerque, NM.

Holman, J.P. 1976. Heat Transfer. McGraw-Hill Book Company, USA, p 98.

Huget, R.G. 1985. The Evaluation and Development of Approximate Schemes for the Finite Volume Method. Doctoral thesis, University of Waterloo, Waterloo, Canada.

Launder, B.E. and Spalding, D.B.1972. Mathematical Models of Turbulence, Academic Press, New York.

Launder, B.E. and Spalding, D.B.1974. The Numerical Computation of Turbulent Flow, Comp. Methods Appl. Mech. Eng., vol.3, p 269..

Lillington, J.N. 1981. A Vector Upstream Differencing Scheme for Problems in Fluid Flow Involving Significant Source Terms in Steady State Linear Systems. International Journal for Numerical Methods in Fluids, 1:3-16.

Linteris G.T. and L. Truett. 1995. Inhibition of Premixed Methane-Air Flames by Halon Alternatives, International Conference on Fire Research and Engineering, Sept. Orlando, FL., pp 153-158.

Raw, M.J. 1985. A New Control-Volume-Based Finite Element Procedure for the Numerical Solution of the Fluid Flow and Scalar Transport Equations. Doctoral Thesis, University of Waterloo, Waterloo, Canada.

Raithby, G.D. 1976. Skew Upstream Differencing Schemes for Problems Involving Fluid Flow. Computational Methods for Applied Mechanical Engineering, 9:153-164.

Rasmussen N. N. K. and A.N. Myken. 1994. The Eddy Dissipation Kinetic (EDK) Model for Turbulent Reacting Flow and Numerical Implementation, 25th Symposium (International) on Combustion, The Combustion Institute, Pittsburgh.

Skaggs, R.R. 2002. Fire Initiation And Sustainment For Interim Armor Vehicle (IAV) Fuel Tanks, U.S. Army Research Laboratory, Aberdeen, Maryland.

Skaggs R.R, N.E. Blackwell and S.J. McCormick. 2002. Advanced Fire Suppression Technologies for FCS. Proceedings of the 23rd Army Science Conference. SAAL-TS Graphical Department, Assistant Secretary of the Army (Acquisition, Logistics and Technology).

Westbrook, C.K. and F.L. Dryer. 1981. Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flame, Comb Sci Techn. Vol. 27, pp 31-43.

Process for Generation of Polygonal Target Models for Infrared Sensor Simulation

Brian Miller
Night Vision Electronic Sensors Directorate
U.S. Army Communications and Electronics Command

Robin Kang
Science Applications International Corporation

I. Abstract

US Army Night Vision and Electronic Sensors Directorate (NVESD) has recently developed a process for generating polygonal target models with attributions suitable for integration with thermal modeling code for use with its infrared sensor simulation Paint the Night (PTN). This paper will describe the process for generating those models from high-resolution Combinatorial Solid Geometry (CSG) models and close-up infrared sensor imagery. It will discuss the method of encoding material attributions into the polygonal data format, SGI Performer™ binary format. It will describe the commercial and custom software tools used in this process. Finally, it will show sample synthetic imagery of targets in scenes generated by PTN with integrated one-dimensional steady-state thermal modeling code.

II. Introduction

In support of its imaging sensor simulation program, the U. S. Army's Communication and Electronics Command (CECOM) Night Vision and Electronic Sensors Directorate (NVESD) has developed a process for generating polygonal vehicle models (henceforth referred to as targets) with attributions suitable for integration with signature modeling codes. The process was developed for use with NVESD's Paint the Night (PTN) Forward Looking Infrared (FLIR) scene simulation software, but the same process can easily be applied to create targets in other wavebands and for other simulations. In fact, with sufficient care in the modeling process, the same models may be developed for use in multiple wavebands.

In the past, NVESD has modeled infrared targets by applying an image of the target's thermal signature, called a texture, directly to a 3-D polygonal model of that target. The target signature is typically a thermal image of the real target collected digitally at close range with a thermal imaging camera. This image is radiometrically calibrated and the image mean is adjusted so that reference points in the image and synthetic scene match. The result is a target with a "real" thermal signature. Signatures are collected at night to minimize the effects of solar loading, heating of the target due to the sun.

In order to generate both day and nighttime scenes with the same targets, a method was developed to add solar loading effects to the target. The brightness of a particular part of the target was increased by a predetermined solar loading factor, based on the location of the sun in the scene as determined by the simulation. This allowed the night target signatures to be extrapolated to day scenes. While effective, this method makes some compromises in accuracy. Since a uniform solar loading factor is applied to an entire model, it does not take into account the differences in materials within the target and the effects of solar loading on internal heat sources.

Two key limitations existed in the previous target models. First is an inability to change model signature based on environmental changes other than sun position, such as ambient temperature. Second is the inability to vary thermal signatures based on the operational state of the target. For example, a tank that has been moving has tracks with a brighter thermal signature while a tank that has been stationary has track signatures closer to that of the armor. Or when the tank has been firing its main gun, the signature of the gun barrel is brighter than when it has been inactive.

One solution to these two limitations would be to collect target signatures for each environmental condition, operational state of interest, and all their permutations. This results in problems. First, collection of all the real imagery for target signatures would be prohibitively time consuming and costly. Second, even if all the required target signatures were collected, dynamic management of the textures within the simulation would provide significant technical challenges, perhaps beyond the capacity of today's hardware.

To reduce the need for data collection, a thermal signature model can calculate the target signature. To get the detail of target signature required for PTN, a very detailed finite element model would be necessary to produce the high fidelity signature images required for a texture. While this solution satisfies our desire to have target signatures at multiple operational states and under varied environmental conditions, both the fidelity requirement of the model and the processing time required to run the model would make this solution practically unworkable for real-time applications.

In order to get the benefits of the calculated model with a manageable level of computer resources we have developed a hybrid solution. This solution utilizes low-resolution thermal modeling and textures based on "real" imagery. A target polygonal model is broken up into thermally similar regions, such as wheels, and attributed with that information. The thermal model can calculate the bulk temperature and radiance characteristics of that region and pass that information on to the simulation that assigns the appropriate color. To achieve greater detail, the image is textured with real imagery, but in this case the imagery is modified to provide only fine scale detail and not the main signature.

This paper will discuss in detail the process of generating target models from high-resolution source geometry and target imagery. It will show some sample synthetic imagery generated by PTN using targets created by this process. This paper will not give details on any specific thermal modeling method or software or discuss the additional information required to implement such a model.

III. Methods

Source Geometry

Most target models used in the PTN simulation originate from high fidelity combinatorial solid geometry models produced by the Survivability Lethality Analysis Directorate (SLAD) at Army Research Labs (ARL). These models were created using ARL's 3-D computer aided drawing package called BRL-CAD™. This package uses a combinatorial solid geometry approach that involves modeling complex 3-D shapes by combining various primitive solid shapes (spheres, cones, cylinders, cubes, etc.) using Boolean operators (union, intersection, difference). In order to use these models in a PTN simulation, the BRL-CAD™ models must be converted into a polygonal format. To address the run-time performance of PTN, the number of polygons in each model must also be drastically reduced.

The BRL-CAD™ model is first converted into the 3-D polygonal geometry format, OBJ, developed by Alias|Wavefront™. This is a popular 3-D polygonal geometry format that organizes the vertex and material data in an easy to follow ASCII file. The conversion is carried out using a command-line executable program distributed with the BRL-CAD™ modeling package. This process retains the hierarchical information of the original BRL-CAD™ model inside the newly created OBJ file. Using a tool developed by NVESD, the OBJ model file is then converted into a SGI Performer™ binary format (PFB), the PTN native format. The resulting PFB target model retains the high fidelity characteristics of the original BRL-CAD™ model. This results in a polygonal model with an excessive number of polygons. At this point, the target models contain on the order of 150,000 polygons

Internal Polygon Removal

A major culprit to the high number of polygons is the detailed interior representation of the target models. PTN is only concerned with the parts of the target that can be seen from the outside. Any interior polygons that cannot be seen by a sensor can be removed. NVESD has developed a tool that effectively removes all internal polygons of the target model leaving only the external shell of polygons. There is usually a 10-fold reduction in the number of polygons once the interior polygons are removed.

Geometry Manipulation with Maya

At this point, the target models still contain an excess number of polygons. A typical target model can be represented by less than 5000 polygons, without significant loss of fidelity from the BRL-CAD™ original. In order to further process the target models and reduce the number of polygons, a software package called Maya is used. Maya is an industry leading,

comprehensive 3-D modeling package developed by Alias|Wavefront™. Maya version 3.0 was used for this effort, but subsequently version 4.0 has been released. It has many powerful features and tools for manipulating geometry and applying textures. The most compelling feature of Maya is the accessibility of its core application programming interface (API). Using the Maya API, developers can create customized tools that can be used within Maya via plug-ins. This feature was immediately put into use to add the capability of importing and exporting PFB geometry files into Maya. A plug-in was developed by NVESD that can bring PFB geometry into Maya along with the hierarchical information and then export a PFB file that retains the hierarchical, texture, and material properties information.

Polygon Reduction in Maya 3.0

Once the target model is brought into Maya, an array of polygon manipulation tools are used to further reduce the number of polygons. There are two main approaches for reducing the polygon count. The first approach is to completely remove objects from the target model that are deemed unnecessary for the proper representation of the models in PTN. These objects are those that do not significantly add to the profile of the target or those that can be represented by the application of textures. The second approach for reducing the number of polygons is to replace complex objects with similar and less polygonal intensive objects. For application in the PTN simulation, it is unnecessary for many objects on the target models to be so highly detailed and represented with so many polygons. These objects include the wheels, headlamps, gun barrels, and hatches. A much simpler model of these objects can usually be created in Maya and used to replace the originals. Figure 1 shows a BMP1 model with the drive, idle, and guide wheel highlighted in black. Combined, these wheels contain over 3,000 polygons. After these wheels are replaced with simpler, polygonal cylinders the number of polygons drops to under 1,000. The overall physical appearance remains the same.

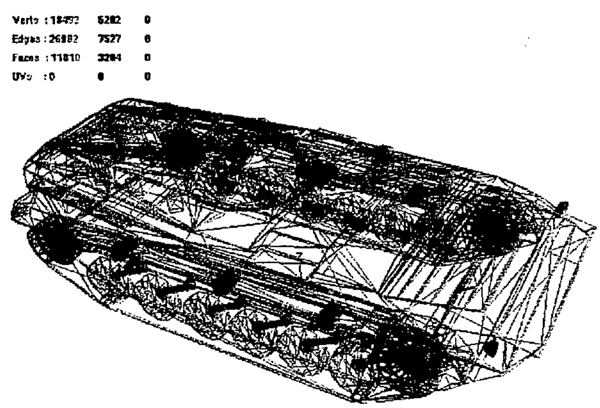


Figure 1 – The drive, idle, and guide wheels contain a combined 3,284 polygons.

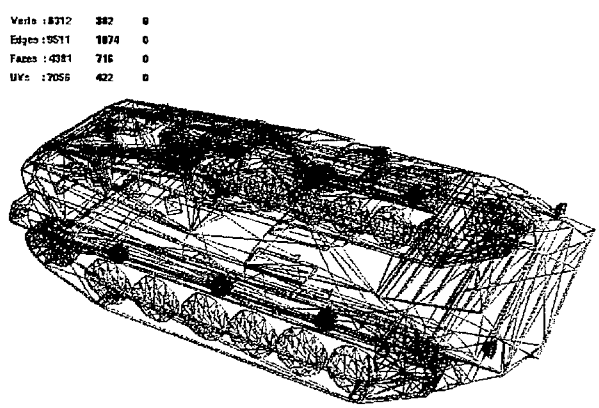


Figure 2 – By replacing the original wheels with simpler polygonal objects the number of polygons goes down to 716.

Because the original source models came from a combinatorial solid geometry format and were not optimized for polygonal representation, even simple objects can consist of an excess number of polygons. For example, one side of a simple rectangular armor plate does not need to be represented by more than two triangles. However, after converting the original source models into a polygonal format, that armor plate may be represented with several times more triangles than necessary. Removing the extraneous polygons and replacing it with fewer triangles provide another way for reducing the number of polygons. Most target models can be represented with fewer than 5000 polygons without significantly impacting the visual characteristics of the models.

Identifying and Regrouping Thermal Regions
The next major step in the PTN target build process is to identify the thermal regions of the target, and to regroup the regions within the target model accordingly. The thermal regions are identified using infrared sensor images taken at various target viewpoints. These infrared sensor images will also serve as textures for the target models. In some cases, the thermal regions closely correlate with a group of physical objects within the model. For example the grouping of wheels may have the same thermal characteristics and therefore represent a single thermal region. In other cases, the thermal regions do not correlate to any single physical object and may be irregular in shape. An example would be the thermal hotspot on the hull of a tank that corresponds to the area over the engine compartment. This thermal region must be carved out of the target model geometry and regrouped as a separate physical region. Figure 3 shows how the thermal regions of the BMP1 were broken out. Each different grayscale region in the Figure 3 represents a distinct thermal region. The hierarchy of the BMP1 target model is shown in Figure 4. Each thermal region is broken out into separate groups. In the case of the BMP1, the geometry was broken out to 12 distinct thermal regions: main hull, front hull, turret, guns, deflector, tracks, glacis, hatches, intake grill, exhaust grill, wheels, and side skirts.

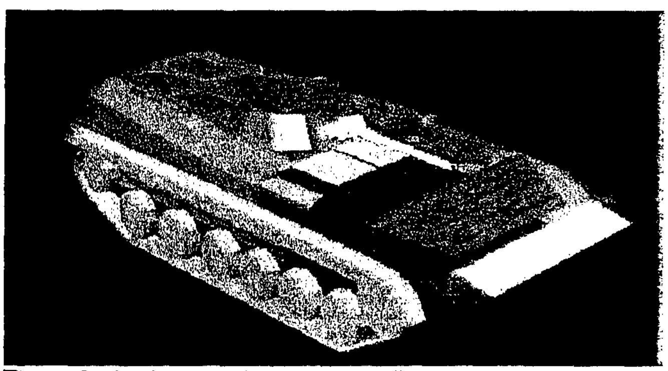


Figure 3 – Each grayscale represents a distinct thermal region.

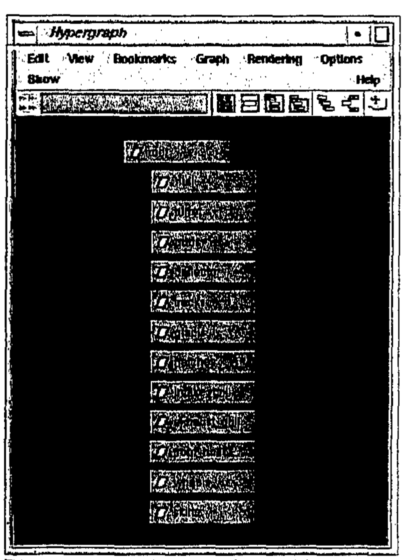


Figure 4 – Each group in the hierarchy represents a distinct thermal region.

Application of Texture and Assignment of Material Shaders
Once the geometry manipulation steps are completed, the target model must be textured with the appropriate infrared sensor imagery. The process of creating textures from source imagery is detailed in another section. When modeling with Maya 3.0, the texture images must have pixel values that are no greater than 8-bits per channel. The latest release of Maya (version 4.0) does provide support for up to 12-bit image files. Using the tools available in Maya 3.0, the target models are textured with the prepared texture image. Special care must be taken to ensure that the thermal regions in the imagery are in alignment with the corresponding physical regions within the target model. Figure 5 demonstrates how the polygons in the target model are mapped to the texture image.

Once texturing is complete, each corresponding thermal region is assigned to a separate material shader that has been given a unique name that identifies the thermal region. The Phong shader must be used in each case because the export plug-in only supports this particular shader. The exporter will insert the unique name given to each shader into the user data field of the material node of the corresponding Performer group node. This is important because the user data field is what the PTN simulation will use to identify each thermal region of the target models so that the material properties for those can be properly modulated.

Figure 5 – This image shows the mapping of polygonal regions of the target model on the texture image.

Exporting to PFB
The completed Maya target model (see Figure 6) is then exported to a Performer binary format (PFB) using the import/export plug-in. Because of the dynamic range requirements of the PTN simulation, it is important to produce target models that are textured with 12-bit image files. A tool was developed at NVESD that creates 12-bit textured models from the directly exported 8-bit textured models by swapping the internal 8-bit texture reference with a reference to an exactly corresponding 12-bit image reference. The end result is the creation of two different versions of the target model in PFB – an 8-bit textured version and a 12-bit textured version.

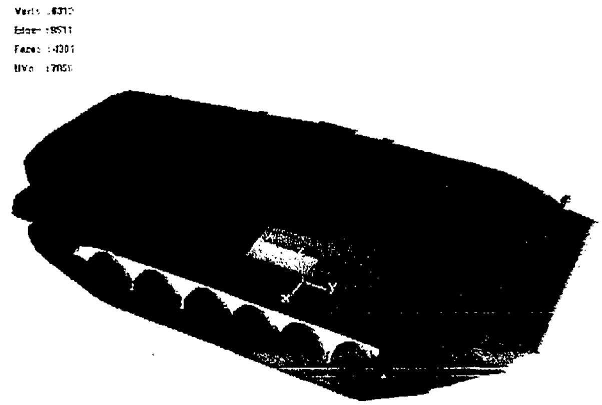


Figure 6 – A finished target model.

Texture Averaging and Thermal Modeling

The thermal regions within each target model will be modulated by using a thermal model that computes and modifies the OpenGL emissivity value of each thermal region during real-time PTN simulations. For this method to work properly, the texture image file must be modified so that the mean value for all pixels within the texture image that are mapped to a particular thermal region must be shifted to some specified mean value. This process must be repeated for each thermal region so that the resulting texture image has corresponding regions that each have the same average pixel values. A tool was developed at NVESD to achieve this goal. This tool iterates through each group corresponding to a thermal region within the target model, averages the pixel values in the texture image mapped to that group, determines the multiplicative factor that the average pixel value is off by (from the specified mean value), and then multiplies each pixel within that region by that factor. This effectively shifts the mean pixel value of each region to the specified value. Figure 7 displays a BMP1 target modeled that has been textured with an averaged texture image.

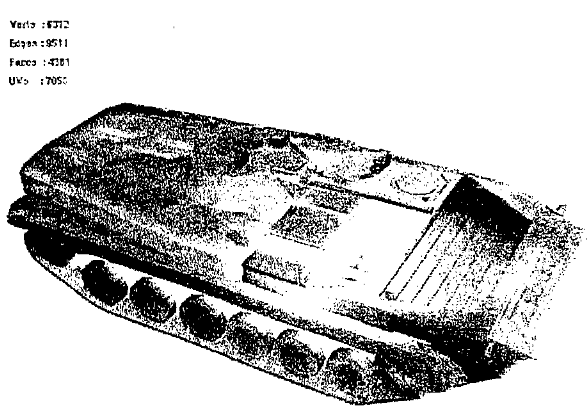


Figure 7 – A target model whose texture has been averaged.

Creation of IR Textures

Textures are created from real imagery of targets taken at night of vehicles in a fully operational state. The imagery has been collected using a calibrated thermal imager. A texture consists of the four principal orthogonal ground views of the target and a top view if available. The processing of the texture imagery is done using ImageJ, a Java based freeware image processing package developed by the National Institutes of Health. The views are read into ImageJ from IMG files using an ImageJ plug-in developed by NVESD. The imagery is then calibrated based on the imager's internal calibration information stored in the IMG file header. Then the means of the images are adjusted so that a reference value in the image matches the same reference value in the simulation. This reference may be a thermal reference source or, if none is available, the

temperature of the ground in the image. This ensures that imagery taken under different conditions matches, and that the imagery matches across targets. The images are then cropped and pasted together into a single image. This image is then resized to make a texture image where the image extents are a power of two.

IV. Sample Imagery

Figure 8 shows multiple synthetic advanced FLIR images generated by Paint the Night (PTN) using an integrated 1-D thermal model, and targets generated using this process. The scene contains a BMP1 and a T72 in day (Figure 8A), night (Figure 8B). and crossover scenes (Figure 8C).

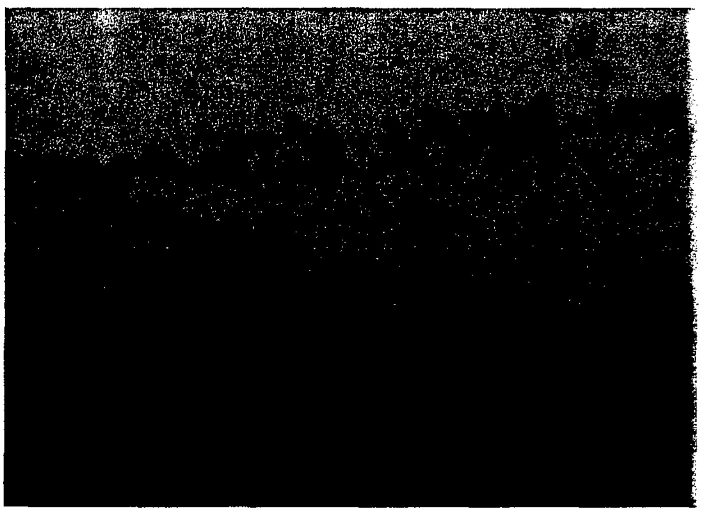


Figure 8A – PTN Scene Daytime (1500 hours)

Figure 8B – PTN Scene Nighttime (2300 hours)

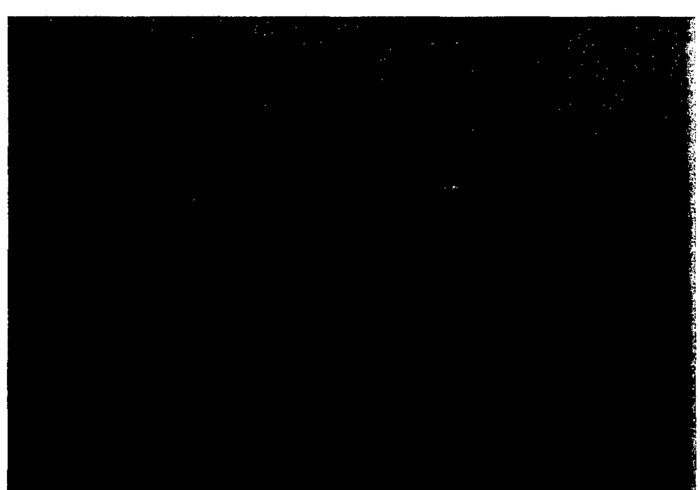


Figure 8C – PTN Scene Morning (0700 hours)

V. Conclusions

This paper described a process for the generation of polygonal targets suitable for use with thermal signature modeling codes. While some coding was required particularly in the area of geometry format conversions and image manipulation, the bulk of the operations were performed on commercial, government owned, or free software. While these models were developed for PTN, the process is not limited to a particular thermal model or type of simulation. In fact, the same process can easily be applied to create targets in other wavebands and for other simulations, and with sufficient care in the modeling process the same models may be developed for use in multiple wavebands.

3D NUMERICAL STUDY
OF
VELOCITY PROFILES AND THERMAL MIXING IN PASSIVE, INFRARED SUPPRESSION DEVICES FOR GAS TURBINE ENGINE DRIVEN GENERATORS

Neal E. Blackwell, Ph.D
Army Power Division
U. S. Army CECOM Research, Development and Engineering Center
Fort Belvoir, VA 22060-5816

ABSTRACT

A 3D numerical study was conducted and compared to experimental data from a venturi-type, infrared (IR) suppression device for exhaust ducting. The results, for a gas turbine driven generator, yield a novel design that is more compact and allows for shorter duct lengths, hence enabling associated camouflage netting to be lower in height and also reducing engine power loss due to reduced backpressure.

Venturi-type, infrared suppression devices are of interest due to low-cost, low-maintenance, duct-type construction. This information is reported to demonstrate the value of virtual prototyping and testing for design optimization, hence impacting system design and providing the best product to the soldier.

INTRODUCTION

Computational Fluid Dynamics (CFD) is becoming a increasingly popular tool for designing air-mixing types of IR suppression devices, for engine exhaust gases, that are integrated into the system design, so that space claim intensive, add-on kits are not required. Venturi-type, infrared suppression devices are of interest due to low-cost, low-maintenance, duct-type construction (Figure 1).

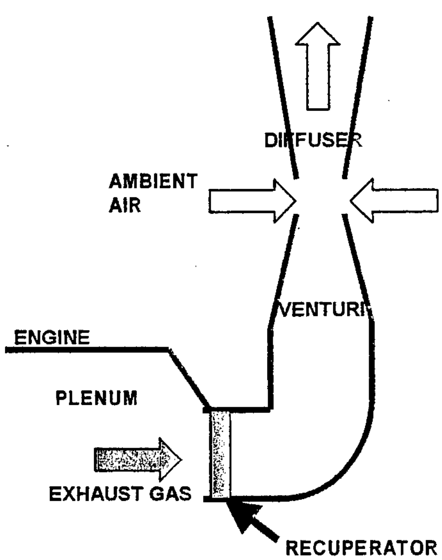


Figure 1. IR Suppression Device with Venturi for Drawing in Cool Ambient Air.

Space claim and engine power loss for venture-type, exhaust IR suppression devices are often the driving factors, over cost and weight, in design selection because venture-type, air-mixing designs are typically inexpensive and lightweight. Judicious use of lightweight and inexpensive ducting materials is common.

Space claim, due to duct length requirements, is typically driven by the length of the ducting required to accelerate the hot exhaust in the converging section of the venturi, so that 1) pressure of the accelerated, hot exhaust drops below atmospheric, hence, drawing in cooler ambient air and 2) the fluid is accelerated at a rate that does not cause excessive exhaust backpressure to the engine. Backpressure must be limited because the performance of the engine used in this study was degraded by backpressure following the equation:

$$\text{Engine power loss (hp)} = 0.367 * \text{backpressure (inches of H20 gage)} - 0.367 \qquad (1)$$

Historically, Army generator sets for this application marginally pass the full-load, high temperature and full-load, high altitude tests without parasitic venturi losses (Howden, 1990 and Monaghan, 1990). Improvements to reduce pressure losses from exhaust driven, IR suppression devices can make the difference in passing and failing these tests. Additionally, engine fuel consumption decreases as exhaust backpressure increases.

Additional ducting is required after the convergence section of the venturi for proper mixing of ambient air with hot exhaust. In addition, this latter section of ducting is often a divergence section to 1) enhance mixing and 2) decelerate the flow to provide longer residence times to cool the mass of the individual particles that radiate to IR sensors. Hence, the main disadvantage of space claim, mainly the duct length, can be minimized if these converging and diverging sections can be integrated into system designs. Typically, the top of these IR suppression devices are designed to be flush with the camouflage netting (Walker, 1991). An integrated design could result in shorter duct lengths, hence enabling associated camouflage netting to be lower in height, perhaps hiding the weapon system more effectively.

This follow-on study was a continuation of the work by Blackwell (1991 and 1992) for the purpose of 1) including the more comprehensive 3-dimensional nature of the accelerating flow, 2) including the solution of the coupled energy equation for thermal mixing, 3) including smooth curved surfaces and 4) demonstrating an example of CFD modeling to converge upon an integrated design that reduces space claim and engine power loss. The CFD model used in Blackwell (1992) 1) was a 2-dimensional model, 2) did not include the energy equation, hence, thermal mixing was not included, and 3) was a single block code that did not have the capability of modeling the duct bends as smooth curved surface.. Rather, Blackwell (1992) used a stair-stepping approach to approximate bends in the duct. The 3D CFD code used in this study is a multi-block code with body-fitted coordinate capability for each block. Hence, the smooth bends of the exhaust geometry were more accurately modeled in this study.

Commercially available gas turbine and diesel engines are often used in military systems to avoid the high cost of development and low quantity production. Hence, such engine designs typically do not include IR suppression of the particles in the exhaust. Even with low particle concentrations in gas turbine exhaust streams, there is still a concern for exhaust IR suppression (Petraska, 1992). This is particularly true for heating secondary objects, such as fine soil particles which become airborne or overhead tree leaves and branches.

METHODS

Numerical Code Description

A numerical, conservative finite volume formulation was used to solve the steady state equations for mass, momentum, energy and the transport equations for turbulent kinetic energy and the dissipation of turbulent kinetic energy. The commercial code is licensed by AEA Technology under the name of CFX4 and has multi-block and body-fitted coordinate capability. For this work, discretation in space was conducted with the 1st order, Hybrid Upwind scheme for advective terms, which, upon convergence, were later switched to the 2nd order, Higher Order Upwind scheme. Also, the 2nd order, Central Differencing scheme was used for the diffusion terms. Velocity – pressure coupling, for continuity, was achieved using the SIMPLEC algorithm (Van Doormal and Raithly,1984) for collocated (non-staggered) grids, which was developed on the basis of the SIMPLE algorithm (Patankar and Spaulding, 1972) for staggered grids. Rhie Chow

Interpolation (Rhie and Chow, 1983) was used to smooth out checkerboard oscillations in pressure and velocity and yield an error on the order of the 4^{th} derivative. Pre-processing was conducted with MeshBuild from AEA Technology and post – processing was conducted with FieldView by Intelligent Light.

Boundary Conditions

Engine exhaust flowed from the manifold, to a large plenum and then through recuperator fins located at the inlet of the exhaust duct (Figure 1) (Blackwell, 1991 and 1992). This arrangement was expected to encourage uniformity of the velocity profile, hence, a uniform velocity profile was assumed. The inlet Reynolds number was $9.7x10^4$ and the exhaust gas temperature at the inlet was specified as Tin from measurements by Blackwell (1991). Turbulence intensity at the inlet was specified as 30% (Martin, 1989). Turbulent kinetic energy and the dissipation of turbulent kinetic energy at the inlet were calculated using the equations by Launder and Spalding (1972 and 1974).

A pressure of 0.0 gage, was specified at the boundaries of the ambient space. Rather than simply applying this boundary condition at the ambient air intake slot, the surrounding ambient space was gridded in order to increase the accuracy of the predictions of air being drawn into the venturi. The ambient space grid allowed for a more accurate model of the shear layer interaction at the ambient intake and also allowed for more accurate predictions of the pressure distribution in the vicinity of the intake. Ambient air was specified as 32 C (90F).

DISCUSSION

Baseline Simulation: M&S Comparison With Measurements

3D CFD simulations agreed well with predictions from Blackwell (1992). No IR suppression devices were installed for these baseline measurements.

Novel Integrated Design To Take Advantage Of Existing Geometry

A novel integrated design (Blackwell, 2002) was possible using the predicted pressure distribution resulting from the 90 degree mitre bend in the exhaust duct. An existing separation region that produces a low pressure could be used to draw in ambient air for IR suppression. An ambient air inlet installed in the low pressure region. This integrated design was modeled to predict the exhaust cooling and the engine power loss due backpressure.

Comparison of Exhaust Gas Temperature Reduction

The novel design reduced the maximum temperature of the engine exhaust gases to a lower value and with a lower space claim or duct length requirement than the venturi case (Figure 2). Assuming that the engine exhaust gas contains soot particles that exhibit black body radiation, the novel design reduces the emissive power of the soot particles by a factor of $T^4_{venturi}$ minus T^4_{novel}. The maximum local pressure drop between ambient air and the venturi ambient air intake was predicted to be a factor of 3.6 times larger than the pressure drop at the ambient air intake of the novel design. The ratio of the novel/venturi total surface area of the ambient intakes was 1.03. However, the mass flowrate of ambient air into the novel design was predicted to be a factor of 1.9 times larger than the ambient air flowrate for the venture case. The relatively low mass flowrate of ambient air in the venturi case is estimated to be due to the resistance of the ambient air velocity component normal to the intake slot due to the relatively large downstream momentum of the venturi jet. It is estimated that the high rate of temperature reduction, of the maximum exhaust gas temperature, in the venture case, was higher than that of the novel design because of higher levels of turbulent kinetic energy (TKE) generated by the venturi jet. The engine backpressure in the venturi case was of factor of 2.6 times larger than for the novel design. As a result, the engine power loss due to backpressure in the venturi case was 3.6 larger than for the novel design case. Hence, the novel design reduced engine power loss by 72%. The challenge with practically applying the novel design is that the

ambient air intake is typically located 1) in close proximity to the engine and 2) inside of a generator housing. As a result, a short, well-insulated intake duct may be required.

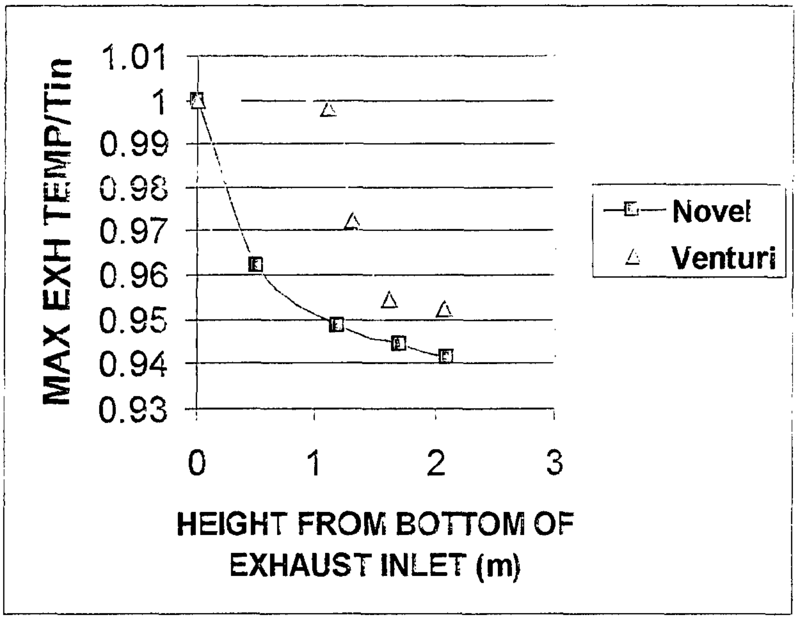


Figure 2. Comparison on Maximum Exhaust Gas Temperature as a Function of Height.

CONCLUSIONS

The novel design reduced the maximum temperature of the engine exhaust gases to a lower value and with a lower space claim or duct length requirement than the venturi case. Typically, the top of these IR suppression devices are designed to be flush with the camouflage netting (Walker, 1991), hence, the lower space claim of the novel design will enable associated camouflage netting to be lower in height, perhaps hiding the weapon system more effectively. Also, the novel design reduced engine power loss by 72% because of reduced exhaust backpressure. Future work will extend the comparisons in this paper to include the predictions of a venturi design by Haase (1990) and thermal mixing enhancement with the counter rotating vortices that were observed by Mattingly and Yeh (1991) in 90 degree bends in circular pipes.

ACKNOWLEDGEMENTS

The author wishes to thank 1) Dr. Grayson Walker, Mr. Rob Hall, Ms. Beth Conway and the former Counter Surveillance and Deception Division at Fort Belvoir for involving me this work, 2) Mr. Kenneth Howden, Mr. William Monaghan Jr. and Mr. Du Wayne Eifert for their valuable insights and 3) my management for allowing me to pursue this idea.

REFERENCES

Blackwell, N.E. 1991. Effects of Velocity Profiles of Turbine Engine Exhaust on Venturi-Type Infrared Suppression Devices. Proceedings of the Second Annual Ground Target Modeling and Validation Conference, Edited by A. M. Lahaie, Keweenaw Research Center, Michigan Technological University, Houghton, Michigan., U.S. Army Tank-Automotive Command Research, Development and Engineering Center, Warren, Michigan. pp 286-302. Limited to DOD and DOD Contractors.

Blackwell, N.E. 1992. Numerical Study of Velocity Profiles in Venturi-Type Infrared Suppression Devices. Proceedings of the Third Annual Ground Target Modeling and Validation Conference, Edited by A. M. Lahaie, Keweenaw Research Center, Michigan Technological University, Houghton, Michigan, U.S. Army Tank-Automotive Command Research, Development and Engineering Center, Warren, Michigan. pp 341-352. Limited to DOD and DOD Contractors.

Blackwell, N.E. 2002. Novel Design for an Integrated, Passive IR Suppression System for Gas Turbine Engine Exhaust Gas. Technical Report, U.S. Army CECOM RD&E Center, Fort Belvoir, VA 22060-5816. Limited to DOD and DOD Contractors.

Haase, Robert. 1990. Faxed Information. Michigan Technological University, Keweenaw Research Center, Houghton, Michigan. Limited to DOD and DOD Contractors.

Howden, K.C. 1990. Personal Communication. U.S. Army Belvoir Research Development and Engineering Center, Environmental Control and Systems Support Division, Fort Belvoir, Virginia.

Launder, B.E. and Spalding, D.B.1972. Mathematical Models of Turbulence, Academic Press, New York.

Launder, B.E. and Spalding, D.B.1974. The Numerical Computation of Turbulent Flow, Comp. Methods Appl. Mech. Eng., vol.3, p 269.

Martin, J.J. 1989. Numerical Solution to Automotive Passenger Compartment Cooling by use of PHOENICS. Masters Thesis, Department of Aerospace Engineering, University of Missouri – Rolla.

Mattingly, G.E. and T.T. Yeh. Effects of Pipe Elbows and Tube Bundles on Selected Types of Flowmeters. Flow Measurment Instrumentation, Butterworth-Heinemann, Ltd., Vol 2, January.

Monaghan, W.C. 1990. Personal Communication. U.S. Army Belvoir Research Development and Engineering Center, Test and Test Policy Division, Fort Belvoir, Virginia.

Patankar. S.V. and D.B. Spaulding. 1972. A Calculation Procedure for Heat, Mass and Momentum Transfer in Three-Dimensional Parabolic Flows, Int. J. Heat Mass Transfer, vol. 15 p. 1787.

Petraska, J.W. and J. A. Guinn. 1992a. Modeling the Infrared Signature of Engine Exhaust Plumes, Proceedings of the Third Annual Ground Target Modeling and Validation Conference, U.S. Army Tank-Automotive Command RD&E Center, Warren, Michigan, pp 331 – 340. Limited to DOD and DOD Contractors.

Rhie C. M. and W. L. Chow. 1983. Numerical Study of the Turbulent Flow Past an Airfoil with Trailing Edge Separation, AIAA J1, 21pp. 1527 – 1532.

Van Doormal J. P. and G. D. Raithly.1984. Enhancements of the SIMPLE Method for Predicting Incompressible Fluid Flow, Numerical Heat Transfer, 7pp.147-163.

Walker, G. W. 1991. Personal Communication. U. S. Army Belvoir Research Development and Engineering Center, Counter Surveillance and Deception Division, Fort Belvoir, Virginia.

Suppression of Thermal Emission from Exhaust Components Using an Integrated Approach

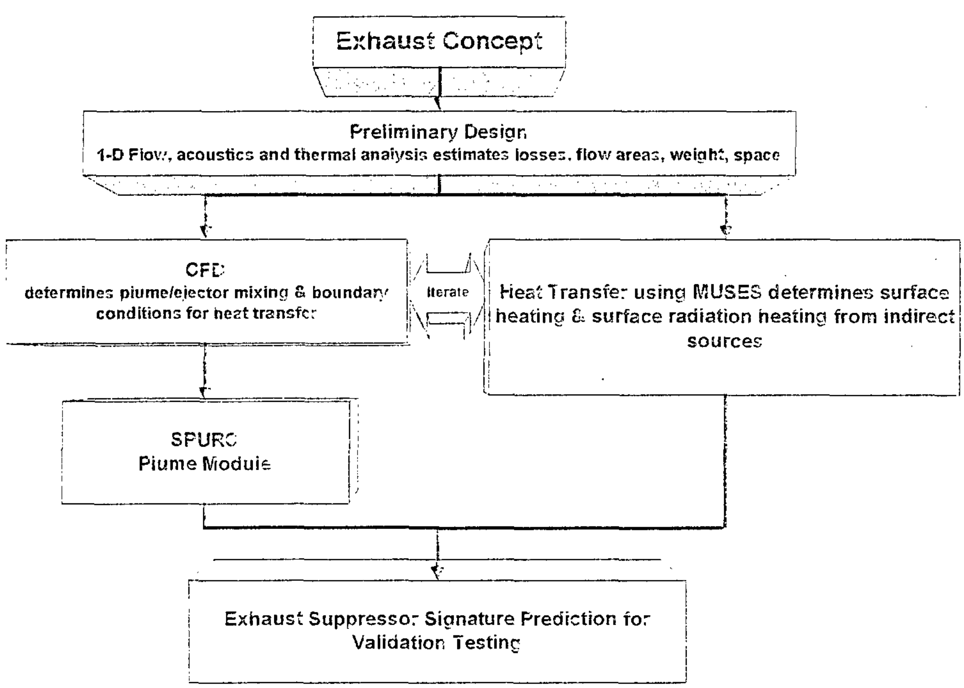
James L. Means, Lance C. Grace
Temeku Technologies Inc.
Lutz, Florida 33548-07222 USA

ABSTRACT

This paper provides the results of an Phase I SBIR to develop an exhaust suppression virtual design model. and exhaust suppression optimization using modeling techniques and hot flow testing. The virtual design model provides an estimate of space required for the exhaust, backpressure to the engine, system weight, gas temperature distributions, and the interaction of the plume with external surfaces both on the vehicle and the ground, and the thermal interaction of the exhaust system with external surfaces through internal convection, conduction, and radiation. A hot flow exhaust test rig has been designed that simulates both exhaust flow and engine bay cooling air flow. This is accomplished with two separate gas producers mounted on flat beds so that the rig is mobile and the sources can be positioned independent of each other.

Virtual Design Model

The virtual design model (VDM) must encompass the disciplines relevant to the performance of an exhaust system. The exhaust system is defined as the hardware necessary to vent the exhaust from the vehicle beginning at the exhaust plane defined by the engine manufacturer and necessary to isolate the exhaust thermally from vehicle structures. The virtual design model must, as a minimum, include an accurate estimate of space required for the exhaust, backpressure to the engine, system weight, gas species distributions, gas temperature distributions, the interaction of the plume with external surfaces both on the vehicle and the ground, and the thermal interaction of the exhaust system with external surfaces through internal convection, conduction, and radiation. Figure 1 depicts the schematic for the virtual design model.



• **Figure 1: Multi-spectral Exhaust System Virtual Design Model**

The modules of the VDM will vary depending on the user and his organization. The VDM should serve as a guide to the use of those modules and provide the input to each module. The VDM must help the exhaust designer interface with CAD

models, structural analysis, ballistic analysis, reliability models, etc. This effort will identify the steps to design and test an exhaust system, sources of data. interfaces with standard design processes, and the degree to which this system is automated.

Computational Fluid Dynamics

CFD analysis of exhaust systems can take weeks to converge and are very difficult to model since the flow regimes include high subsonic compressible internal flow, low velocity wake regions, and regions with high vorticity. To be useful in the design process, the CFD analysis must be responsive to the design process. This can be accomplished by selecting an appropriate CFD code and using it on a subset of the full exhaust system to screen designs, e.g. centerline analysis for high aspect ratio 2-d nozzles.

There are many CFD codes available commercially and in the public domain. The emphasis in this program is to select one of the public domain codes to incorporate in the exhaust design model. Of these codes CFL3D, VULCAN, FUN3D, HEFSS, and WIND were selected as the most viable. CFL3D is a public domain CFD program maintained by NASA Langley Research Center. CFL3D has been used extensively by the U. S. Naval Surface Warfare Center, Carderock Division, and NSWCCD for predicting the flow in exhaust systems and over ship superstructure. All of the commercial codes are available for application on Windows NT. CFL3D is available only for UNIX workstations and supercomputers. However, the FORTRAN code is available and can be hosted on an NT based personal computer.

The future for CFD codes are codes using unstructured grids like FUN3D. NASA representatives feel that good results have been obtained in both wakes and wall regions using the 1998 k-ω turbulence formulation. The flow equations will be stiff and convergence will be difficult. The solution will be strongly dependent on the grid. Results of other CFD studies indicate that the solution will also be very sensitive to the specification of the initialization of the variables in the grid. This agrees with results previously obtained. The emphasis is consolidating existing codes into an advanced CFD code. This is High Energy Flow Solver Synthesis, HEFSS. At the present time there are two codes available and support by the Aerodynamic and Acoustics Methods Branch: VULCAN and CFL3D. Both codes will run on a PC if the PC is running LINUX or another version of UNIX. The Fortran code is available and can be ported onto Windows. VULCAN can compute results with chemical species.

CFL3D has been selected as the CFD code for initial use in the VDM since it is a mature code.

Heat Transfer

The heat transfer analysis must include: Radiation between components; Non-isotropic properties; Adequate bi-directional reflectance (BDRF) model; Interaction with foreground/background; Internal radiation and convection; Compatible with advanced treatments; Secondary Heating due to plume; Run time compatible with design process.

Of the many heat transfer and radiation models available, MUSES is far and away the best. It is already being interfaced with CFD codes to get better local convection coefficients.

Plume Analysis

The plume analysis must include: Multiple exhaust plumes; Non-axisymmetric plumes; Specie distributions for diesel/turboshaft engines; Cross-flow; Interaction with ground plane. The only real alternative for plume signature prediction is the adaptation of SIRRM III using the flow field predicted by the CFD program.

Hot Flow Test Rig

The Phase II SBIR will provide a ground vehicle exhaust hot flow test rig that is capable of testing exhaust systems for ground vehicles with engines (up to 700 hp). This plan will be used to identify the site for the test rig, safety and operational issues, sources of components for the test rig, and serve as a basis for estimating the cost of the test rig.

The criteria for a multi-spectral exhaust system test rig include: Ability to control temperature, pressure, mass flow rate independently ; Isolation of diverted flow from exhaust flow; Simulation of External flow for the real installation; Far enough

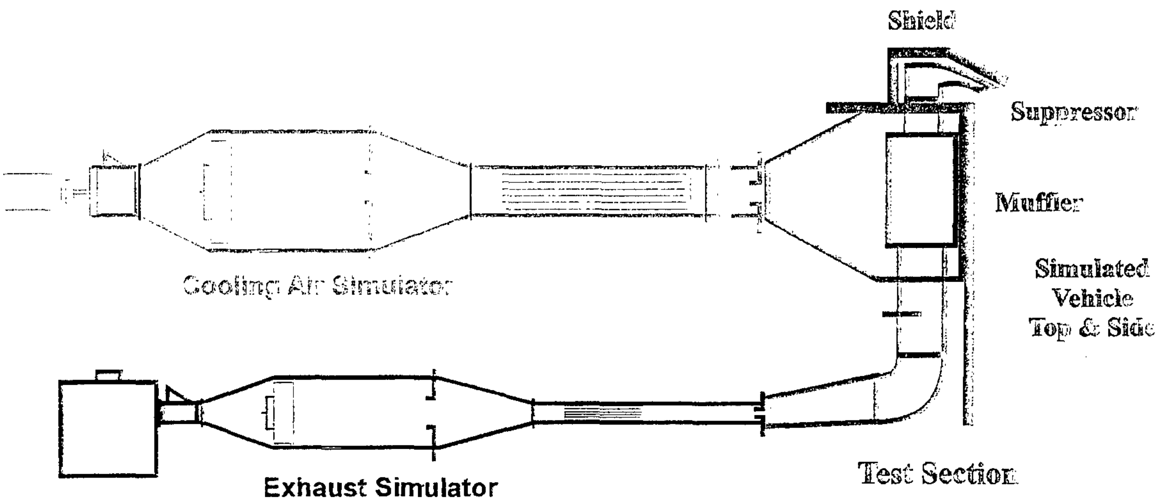
from walls and obstacles to insure that the flow at the test section is not influenced; Compliant with community noise and pollution laws.

The test rig must be able to simulate the full operating range of the exhaust system from idle to full power and must accommodate combinations of engine off-design performance. This is best accomplished when the heat source, the mass flow rate, and the exhaust temperature are controlled independently. This can be accomplished by using a fan with excess flow rate capacity at a backpressure corresponding to the test rig flow loss. Excess mass flow is vented through a waste gate. The waste gate helps obtain the fan maximum pressure by approaching the operating point of the fan from the unstalled side of the fan map. If an engine is used to provide the flow rate and temperature, it must either be the engine for the exhaust system or is capable of providing more than the required mass flow rate and enough gas temperature margins so that the flow can be adjusted to the exhaust system design point.

The optimal exhaust system will function through the full range of flow from idle through max power. Maximum power will correspond to a vehicle operating condition. The optimal test rig will simulate the vehicle installation and the external flow conditions corresponding to the engine power setting. Since most ground vehicles spend a great deal of time at idle, it is also important to be able to simulate cross winds so that secondary heating of vehicle surfaces can be investigated.

Rig Concept

The test rig will have two independent air sources. One air source is required to simulate engine exhaust with enough pressure rise and mass flow rate to simulate engine exhaust at the manifold and the extra pressure drop required by the rig components. Another air source is required to simulate the engine compartment cooling air with enough pressure rise, temperature rise and mass flow rate to simulate the engine compartment pressure drop and rig component pressure drop. Figure 2 depicts this concept.



• **Figure 2: Rig Concept**

The exhaust simulator is designed to provide up to 3,100 standardized cubic feet per minute (SCFM) of air at 60 inches of water pressure at 1,100° Fahrenheit. The engine cooling compartment simulator is designed to provide up to 25,000 SCFM at 12 inches of water pressure at 250° Fahrenheit.

Both systems are very similar. A variable amount of airflow at varying amounts of pressure is input at the left side of the figure. A controllable dump door allows fine-tuning of the air mass and air pressures. A diverging transition section presents the airflow to the combustion section where a propane burner increases the temperature to the desired level. A pitot tube and thermocouple array in the combustion chamber measures conditions that are compared to the measurements from a similar array downstream in the exit tube to provide accurate mass flow rates. A converging transition section smoothly forces the airflow into the exit tube. The air flows through an air straightener before it flows past the second measurement array. The air

exits the right side where extension sections and test item specific transition pieces present the controlled airflows to the test items.

Suppression Techniques

The suppression techniques used for the exhaust system must accommodate suppression in infrared, radar, visual and acoustic spectra. The techniques used in each spectrum must be compatible with the solutions in the other spectra. The Virtual Design Model must be able to analyze and assist in the design the full spectrum of suppression techniques. The Test Rig must be capable of testing and evaluating all of the developmental issues with each of the suppression techniques.

Infrared Suppression Techniques

The techniques used to suppress exhaust signatures are:

- Plume Signature Suppression: Ejector dilution; Fan augmented dilution; Multiple Exhaust plumes with self-absorption and external mixing; Swirl Augmented dilution; Clean exhaust gases
- Hot Part Signature Suppression: Thick Film Cooling; Thin Film Cooling; Fin cooling; Hidden Sacrifice Surface using Coanda Effect; View hiding with insulation; Emissivity Control

Plume Suppression Techniques

Most exhaust systems on diesels and turboshaft engines have exhaust gas temperatures between 850° F and 1000° F. Dilution ratios of 110% to 150% are usually required to meet the desired plume signature. The most direct way to dilute the flow is to use an ejector or a fan to pump cool air into the flow. To insure that the flow is mixed before the exit plane a mixing tube 10 to 20 characteristic dimensions long must be provided in the design. A fan system will require space for either a hydraulic drive, an electric motor, or a mechanical drive from the engine. The power required for the fan system is usually more than the power lost to back pressure for an ejector system. A low loss flow path to ambient air with acceptable radar, visual, and acoustic signatures must be provided for either the fan or ejector to function properly.

An alternative or a complement to the fan or ejector dilution scheme is to use external mixing. External mixing can be promoted to limit the extent of the hot core and provide an area average radiance that is acceptable. There are several ways to promote external mixing. One of the easiest is to break the exhaust flow into multiple jets. This increases the surface area available for mixing, spacing between the jets helps maximize the flow induced for mixing with the jets, and the cool air layers between the jets causes self absorption which reduces the radiance to the seeker. Usually the core temperature of the jet dissipates in 10 to 15 characteristic dimensions from the exhaust plane and nearly completely dissipated in 50-100 characteristic dimensions. Consequently, reducing the characteristic dimension reduces the extent of the hot plume. Swirl also reduces the extent of the plume and enhances external mixing by increasing the path length of the flow and increasing the diffusion of the jet. Combustor designers have long used this principle to maximize mixing in short axial distances.

The primary source of plume radiation is the particulate in the exhaust. Advanced mufflers using catalytic surfaces are available to convert the particulate and CO into CO_2 where there is significant atmospheric absorption.

Hot Part Signature Suppression

Hot part signature suppression can be handled directly by cooling the walls with thick film cooling, thin film cooling or fin cooling to a secondary fluid. Thin film cooling schemes introduce a thin film along the wall and replenish the film before the mixing layer reaches the wall. This introduces a complex structure with many edges to scatter radar energy and little room for radar absorbing material. Thin film cooling is also very susceptible to mal-distribution of flow caused by cross flow resulting in local hot spots. Thick film devices introduce a large flow in a slot that persists for long wall lengths. These devices can be unstable causing transient heating on viewable surfaces. Fin cooling to a secondary fluid attempts to maintain an acceptable wall temperature by increasing the heat transfer rate to such an extent that the difference between wall temperature and gas temperature is maximized.

A second class of hot metal suppressors hides a direct view of hot parts from the threat sensor. This can be used in combination with a thick film cooling/dilution system while using a hidden surface to stabilize the flow and the Coanda effect to turn the flow in a desired direction at the exit plane. When view hiding is used, care must be given to using the

appropriate insulation to prevent conduction or radiation heating of viewable surfaces. Cool surfaces may reflect radiation from hot parts to the threat sensor. If the emissivity of that surface is 1, there will be no reflection. But parts with an emissivity below 1 can have unacceptable reflection of internal hot parts.

Validation Concept

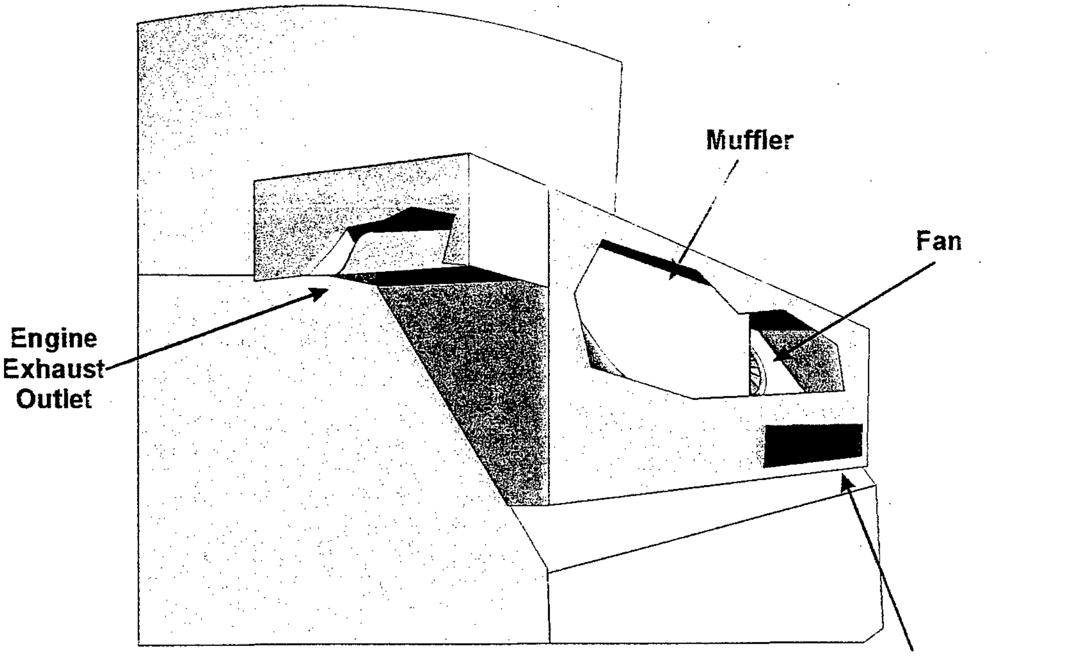
The design to be used as a test case should incorporate:

- Simple Exhaust using Coanda Effect to direct flow away from the vehicle;
- Segmentation into multiple jets to test self absorption;
- Vortex generators to test external mixing;
- Thermal Isolation of Outside surfaces with insulation:
- Redesign of muffler to include removal of particulate while suppressing noise;
- Induced flow to aid in internal heat rejection;
- Structured surface and RAM for other spectra;

The validation design will have three variants to test different aspects of the VDM and the viability of the Hot Flow Test Rig.

The multi-spectral exhaust system validation concept, Figure 3, is composed of an external sponson constructed over the existing facet on a ground vehicle. The engine cooling exhaust is relocated to the sponson where it exhausts out of a LO exit. Cooling air is induced from the inboard interface with the hull and exhausted above the upper lip of the exhaust to prevent the lip of the duct from heating. Insulation with a low emissivity and radar absorbing material will be applied to the interior fan transition duct. The cavity between the transition duct and the sponson will need a radar absorbing material to make any radar return from the cavity negligible. If possible, it is desirable to use a radar absorbing material that acts as insulation and has good acoustic suppression attributes. Cooling air for the lower surface, which is visible to the threat, will be drawn through a slot below the sponson at the interface with the hull. These surfaces will require shaping and radar absorbing material in addition to a structured surface to prevent any reflections from the hot parts above. The transition duct will be attached to the engine exhaust duct before the muffler.

Temeku has developed a concept to insure adequate jet penetration velocity in the engine exhaust at low power settings, the variable flow control. The variable flow control is a blow in door on the exhaust duct. The position of the blow-in door is determined by the relative static pressure between the engine exhaust flow and the fan exhaust flow. The door will weather vane to a position where the two static pressures are equal. When the fan exhaust cannot balance the static pressure of the engine exhaust flow, the door closes and seals. At low engine exhaust flow rates typical of idle, the door weather vanes to equilibrium position and allows fan exhaust air to dilute the flow. Since the cooling fan usually does not have a great deal of back pressure, a low flow loss path must be placed above the muffler to a position just before the exhaust exit so that the fan will be able to pump dilution air. The cooling air exhaust areas will be chosen to permit adequate fan flow to provide the required cooling for the engine while maintaining a high enough pressure in the transition duct to operate the blow-in door.



• **Figure 3: Multi-Spectral Exhaust System Validation Concept**

The first variant will be a simple duct and advanced muffler with advanced insulation. The duct will turn the flow to the horizontal with ejector mixing near the exit. Radar absorbing material will be simulated and the internal geometry will be designed to result in multiple bounces in the radar absorber. A source of ambient air will be provided to the rear surface. A second variant will incorporate the advanced muffler and advanced insulation, but divide the exhaust duct into three separate jets to test the effect of self-absorption and plume interactions. The third variant will incorporate vortex generators in each of the three ducts to test the impact of exit vorticity and the ability of the VDM to predict the flow field.

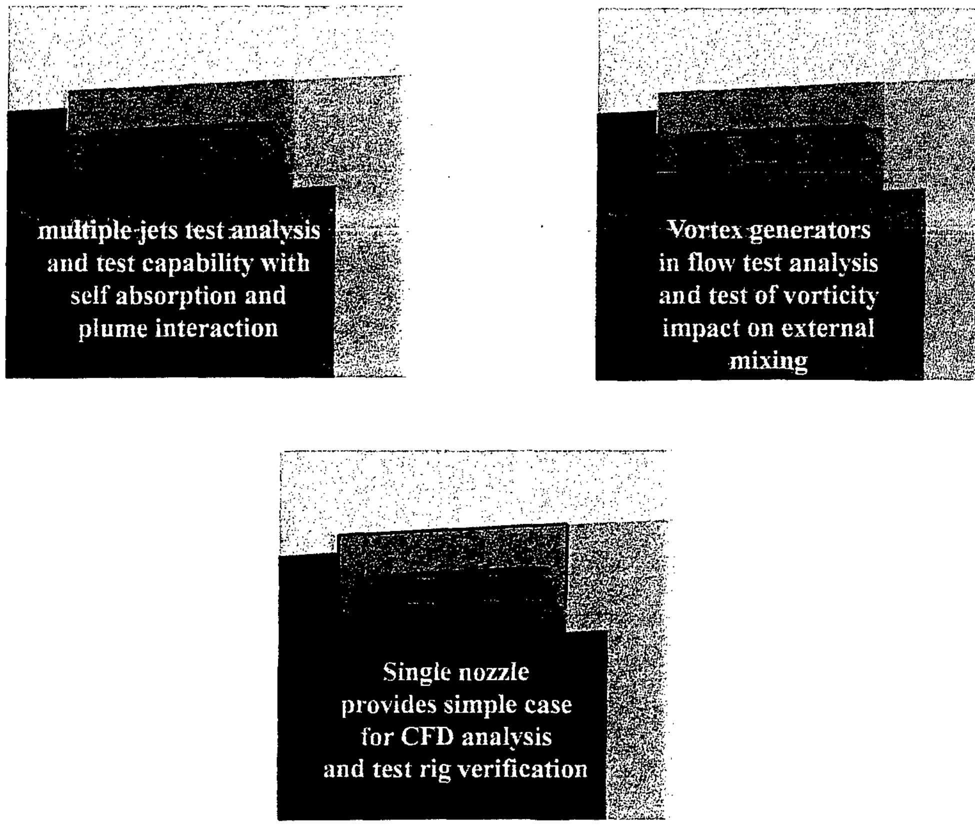
Three exhaust configurations will be tested, Figure 4, during the Phase II effort. The end of the vehicle and the aft portion of the simulated sponson will be simulated on the test section. The simple exhaust without the blow-in door was measured during the Phase I option. During phase II, the complete exhaust system will be simulated with the simple exhaust with the variable flow control, the three nozzle exhaust and the three nozzle exhaust with vortex generators.

The simple nozzle with the blow in door will test the analytical capability to predict the mixing and temperature profiles at the exit. It will test the ability of the test rig to control both flow systems and isolate the control of the two flow systems when they are coupled.

The three nozzle exit will test the analytical capability to predict and optimize induced cold air layers between multiple nozzles and establish the impact of self-absorption on infrared signature. This design will test the ability of the test rig to simulate significant secondary flow paths with isolation from the test rig inlet and waste gate air. It will also test the ability to the rig to measure and document external mixing.

The third concept adds vortex generators to the three-nozzle exit. This will test the analytical capability of predicting external mixing with significant vorticity. Since this exhaust concept spreads radially due to the centrifugal forces on the

exhaust jet, it will stress the test rig's ability to simulate the ground plane with the appropriate thermal characteristics for ground spot analysis.



• **Figure 4: Exhaust duct configurations for analysis, test and validation.**

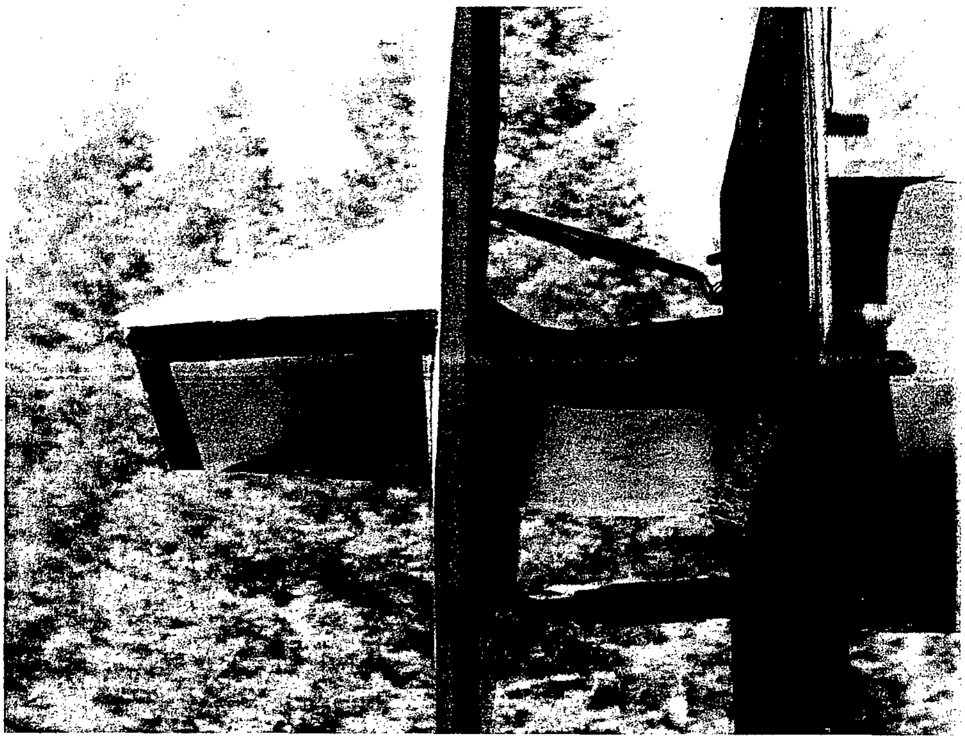
Validation Concept Preliminary Design

Temeku selected a generic design for a typical infrared suppressor for the test nozzle. It was covered by a generic shaped hood. Factors considered for the shapes of both pieces have been previously discussed. The nozzle was fabricated from 1/16 inch thick steel. The hood and simulated vehicle surface were built of wood.

Test Item Description

Temeku selected a generic design for a typical infrared suppressor for the test nozzles. The nozzles were covered by generic shaped hoods. Three separate nozzle configurations were built and tested. Each nozzle had the same side dimensions. In addition, each nozzle had the same cross-sectional throat area of 13 ½ square inches.

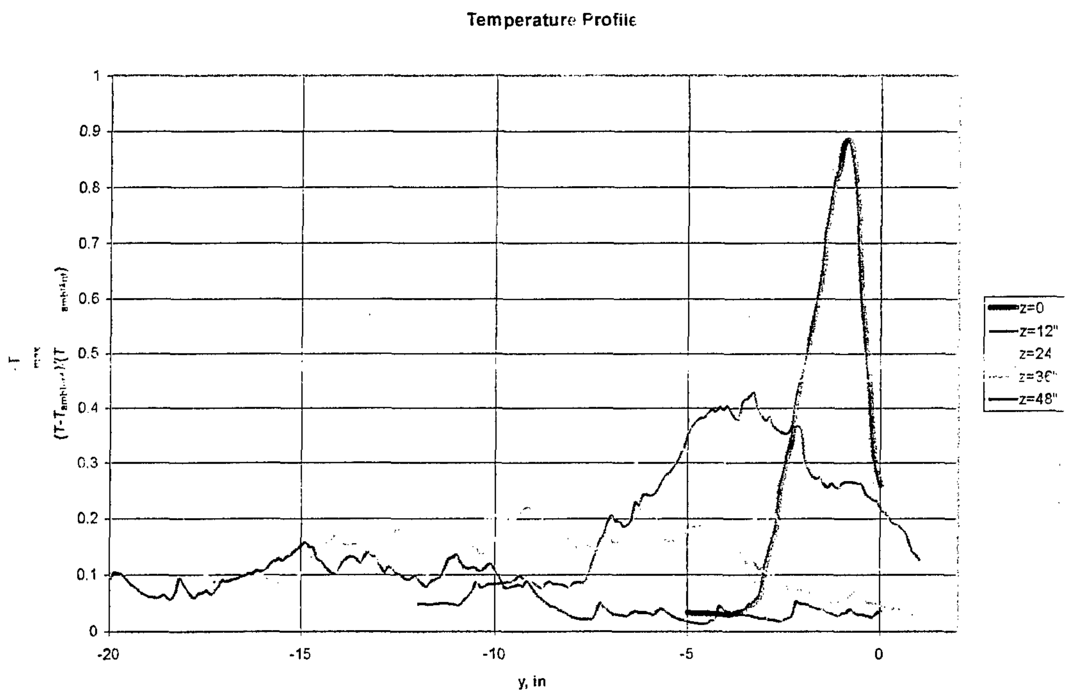
The first nozzle was a single nozzle 9 inches wide. The second nozzle consisted of three 3 inch wide nozzles separated by 3 inch spaces. The third nozzle was the same as the second nozzle but with 4 vortex generators positioned at the center of each wall at the base of the nozzle. The vortex generators in the center nozzle caused clockwise airflow as viewed from outside while the vortex generators in the outside nozzles caused counter-clockwise airflow.



• **Figure 5: Single Nozzle w/ Hood from Side**

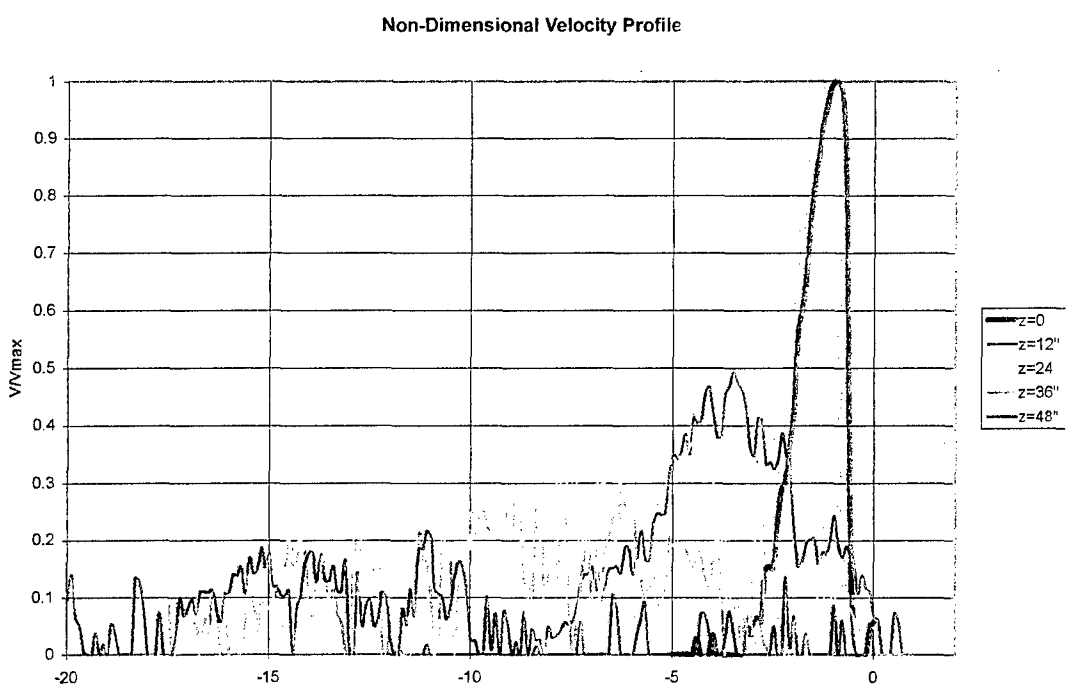
Phase I Test Results

All nozzles were tested at a mass flow rate of 0.9 lb/sec at 680° F. The single nozzle results along the center line indicated that the core flow had already mixed at the nozzle exit and the hot flow was essentially dissipated 36" (12 characteristic dimensions) from the exit.



Page 1 of 1

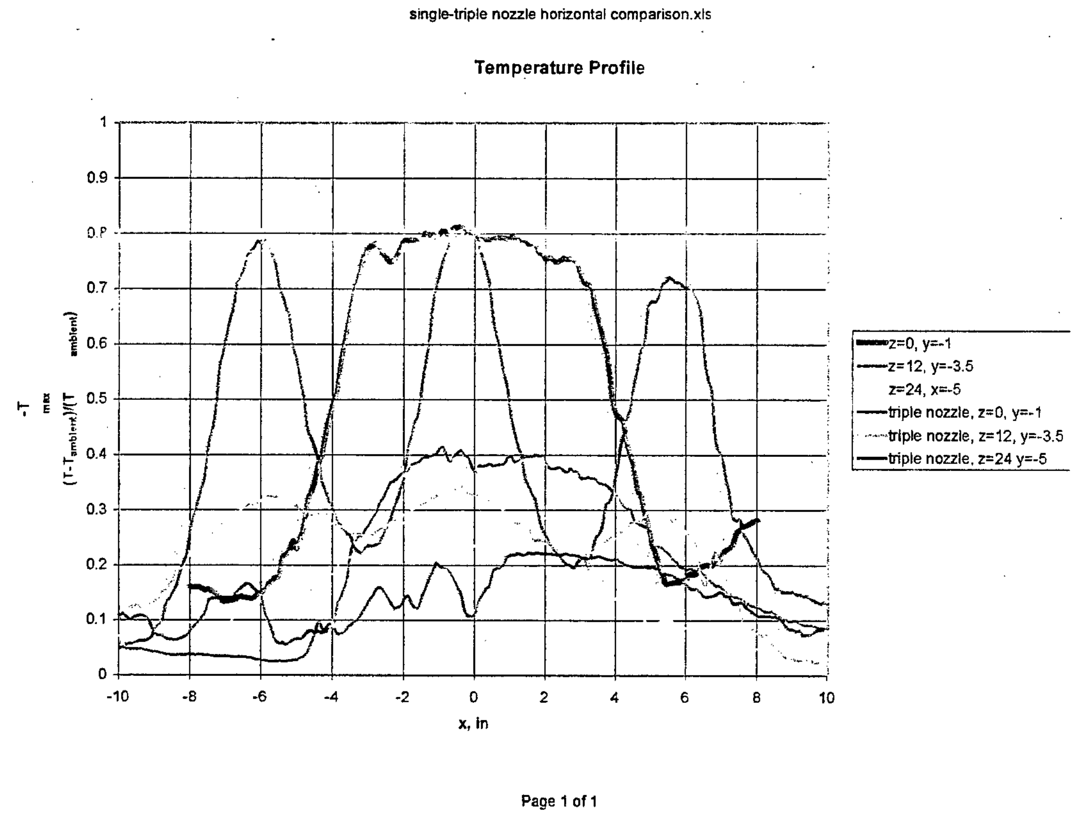
• **Figure 6: Temperature Variation with Distance from the Exhaust for the Single Nozzle**



Page 1 of 1

• **Figure 7: Velocity Profile Variation with Distance from the Exhaust for the Single Nozzle**

The triple nozzle accelerated the dissipation process and created hot streaks that reduce the effective infrared signature of the exhaust by introducing cool air between the hot streaks. This must be evaluated for specific infrared seekers to determine the impact of the cool air on the size of the seeker pixels that average the radiance.



• **Figure 8: Comparison of Triple Nozzle with Triple Nozzle Temperature Profiles**

An examination of the data revealed that the thermocouple on the survey probe responded very quickly to temperature changes while immersed in the hot exhaust flow, but required several seconds to reach ambient temperature when removed from the hot flow. All temperature profiles indicated a region of warm air above the plume. This is an artifact of the time constant of the thermocouple probe. This can be corrected easily in Phase II by adjusting procedures to either stop at each data point or let the thermocouple settle or by taking data from either side of the plume.

Acknowledgements

This SBIR is under the direction of Wally Mick, TACOM under contract DAAE07-02-C-L009. The cooperation and coordination with ThermoAnalytics and Teresa Gonda has been invaluable to structuring a successful exhaust signature design model. Dr. Chris Rumsey, Dr. Robert T. Biedron, and Jeffery A. White at NASA Langley Research Center have contributed to the selection of the CFD program. Dr. T. C. Tai at the Naval Surface Warfare Center has contributed to the use of CFL3D.

An Automated System for Observer Probability of Detection Experiments

Robert Mantey
U.S. Army TACOM
6501 E. 11 Mile Road
AMSTA-TR-R. MS 263
Warren, MI 48397-5000

ABSTRACT

The Army is looking at signature management as a way of increasing the survivability of combat vehicles for upcoming programs such as the Future Combat System. As part of the Army's work in this area, the TACOM Survivability Technology Area has the mission of integrating and maturing advanced survivability technologies for combat vehicle systems. In order to determine the real world effectiveness of new camouflage technologies, it is sometimes necessary to do Observer Probability of Detection (Pd) experiments on military targets. In these experiments multiple military observers are asked to search for a military target positioned within a field of regard on a test range. When a target is detected the time of the detection is logged, and then it is verified whether the target is an actual target of interest or a false detection. Early Pd tests used paper log sheets, range maps, stopwatches, and a variety of other manual devices to complete the experiment. This kind of data collection was very manpower intensive and had many opportunities for errors during the data collection and analysis. In order to reduce the manpower requirements and improve the reliability of the data from Pd experiments, TACOM as developed an automated system for Pd data collection. This new system has a master control system, an instrumented station for each observer, and adds the ability review results immediately after completing a Pd run. This paper presents an overview of this new Pd data collection system and discuses plans for the continued development of the system.

SYSTEM DESCRIPTION

The TACOM developed observer data collection system consists of two trailers with a total of 24 observer stations and instrumentation to automate the data collection. A picture of the observer trailers is shown in figure 1. The point of a field of and observer regard experiment is to determine the Probability of detection (Pd) of a target at different ranges. A Pd test is executed by having military observers search for a target at an unknown location within a defined field of regard. The TACOM developed system has 24 stations with a set of input switches, a digital display and a pan/tilt head instrumented with optical absolute encoders with a sight mounted at each station. There is also a manual controlled electric window shade that is part of the system to block the view of the observers of the range between runs while the target is being positioned. Figure 2 shows an example of a typical Pd test layout:

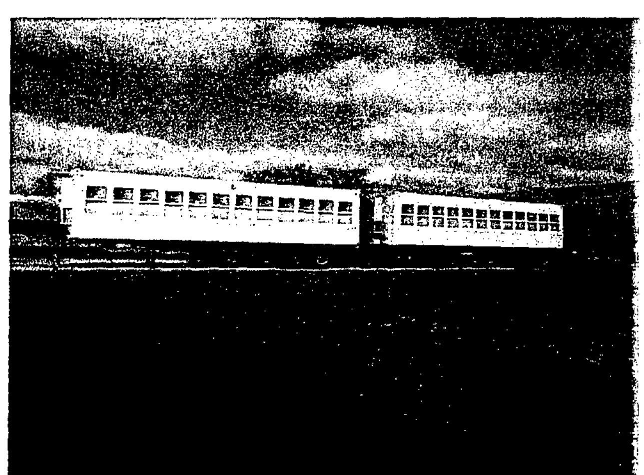


Figure 1 Picture of the Observer Trailers

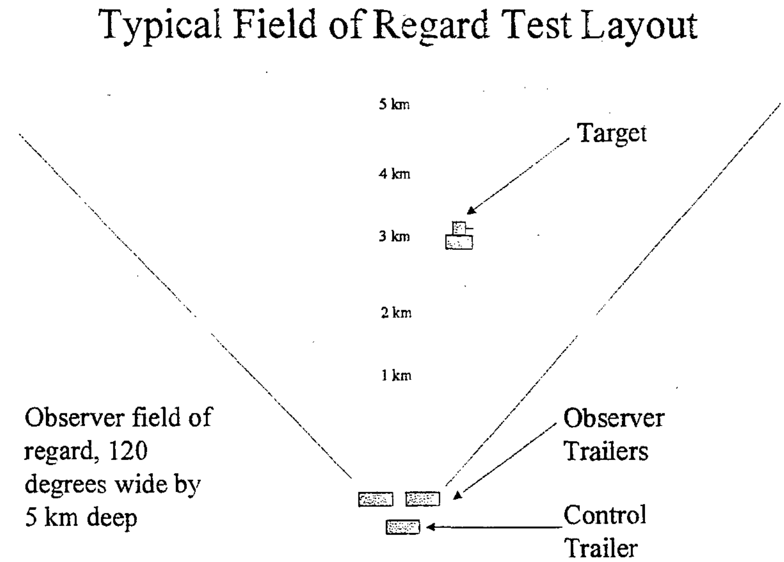


Figure 2 Typical field of regard test layout

A Pd test run is completed as follows: With the window shade down to block the view of the observers, the target is positioned within the field of regard. Then the shade is raised and the observers are given a fixed search time to locate the target, usually two minutes for visual tests. If an observer locates a target, he or she hits the red detect switch, aims the sight on the pan/tilt head on the target, and hits the green verify switch. When the search time has run out, the observers who found a target late in the run are given a few moments to finish verifying the target location and the window shade is lowered again.

Only two sensors have been used on the system to date, they are binoculars and the unaided eye. The standard binocular station has a set of military 7×50 binoculars that the observer holds up to his or her eyes by hand while searching. If the observer finds a target, there is a variable zoom rifle sight that is fixed at 7× magnification with a piece of tape mounted to the pan/tilt head for verification. Unfortunately the rifle sight is designed for about four inches of eye relief which makes it a little tricky to use. A picture of the binocular observer station is shown in figure 3.

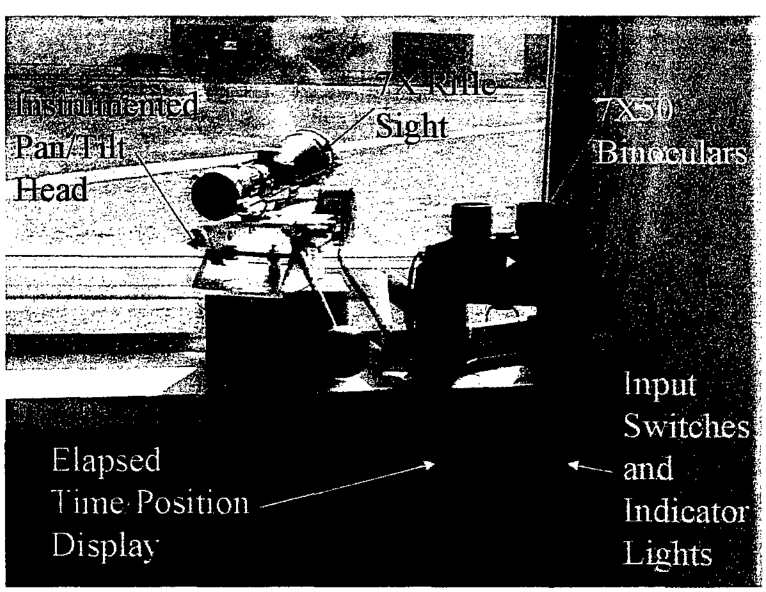


Figure 3 Binocular observer station layout

For the unaided eye station the observer obviously uses his or her naked eye for the search. The unaided eye station has a red dot pistol sight mounted to the pan/tilt head. This sight is nice because it has no eye relief problems and it also corrects for the angle at which the observer looks trough it. If the sight is aimed at a target, any head movement changing the angle of viewing through the sight will cause the red dot projected from an LED to move staying on the target. Figure 4 shows the layout for an unaided eye observer station.

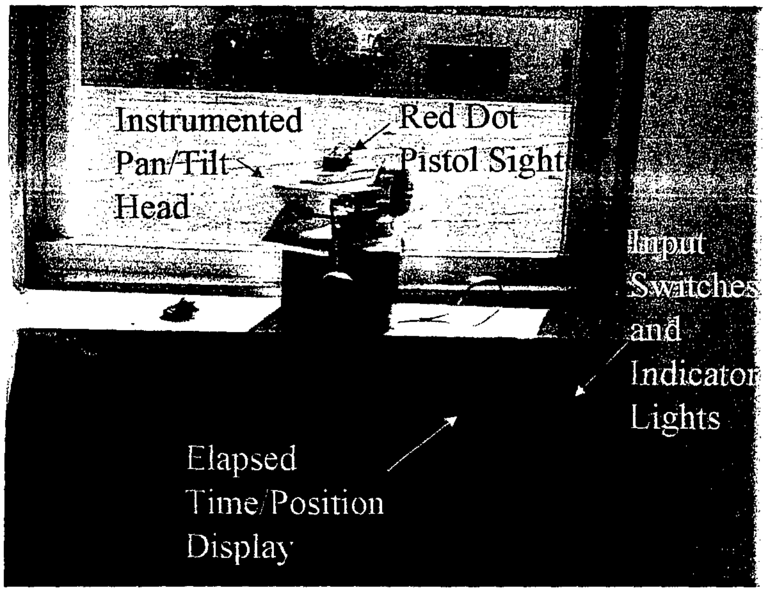


Figure 4 Unaided eye observer station layout

The entire data collection system is computer controlled using National Instruments LabVIEW developed software. LabVIEW is a visual programming language designed for instrument control making it easy to create virtual switches, indicators, displays, controls and communicate with external instruments. A LabVIEW program is referred to as a Virtual Instrument (VI). There is a screen grab of the front panel of the Observer Control System VI in figure 5. The top section of the panel has virtual buttons that control the operation of the system. The middle section lists the detailed information about the run including the range, position, orientation, and timing information. Finally the bottom section of the panel shows the state of all of the observer stations and the Pd for each type of sensor computed in real time during the run. At the end of each run all of the data displayed on the LabVIEW front panel is saved to a tab delimited file that can be easily exported into Microsoft Excel for further analysis.

There is a Master Targeting Station that determines the position of the target within the field of regard. The Master Targeting Station uses a video camera for visual tests and is mounted to a QuickSet precision pan/tilt head on a tripod positioned between the observer trailers. The QuickSet pan/tilt head is remotely controlled and is connected the control PC via a RS-232 port. Its position is continuously monitored while the system is running. When an observer detects a target, the position of the Master Station is compared with the observer target position when the target is verified to determine if the observer has found a real target. Also documentation video of the target during the run can be collected using the Master Station.

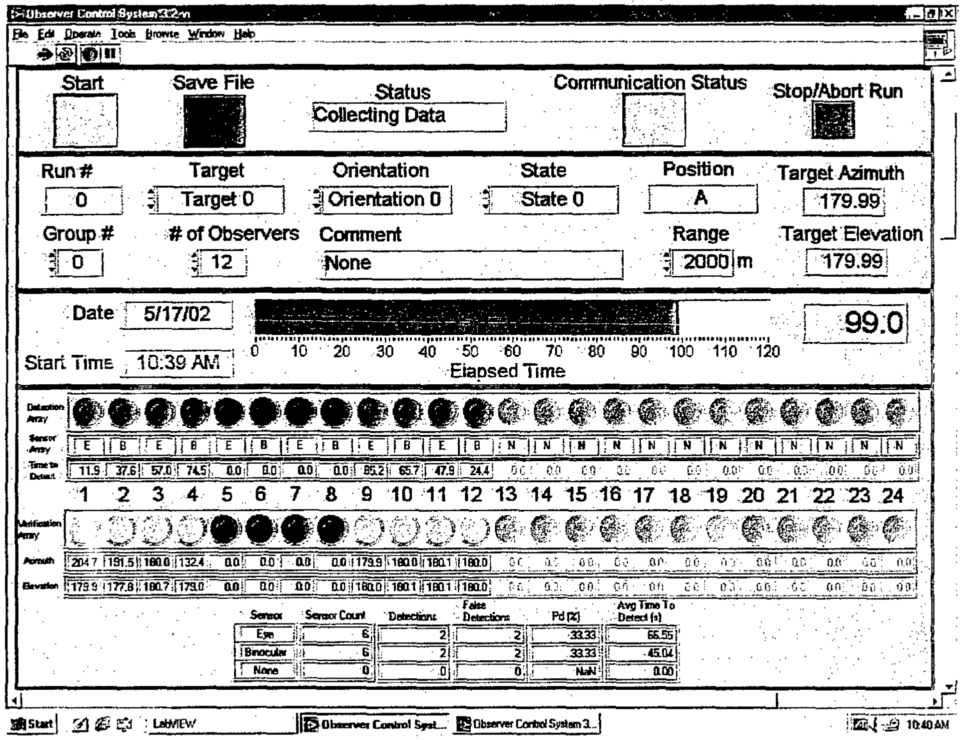


Figure 5 Screen grab of Observer Control System LabVIEW front panel

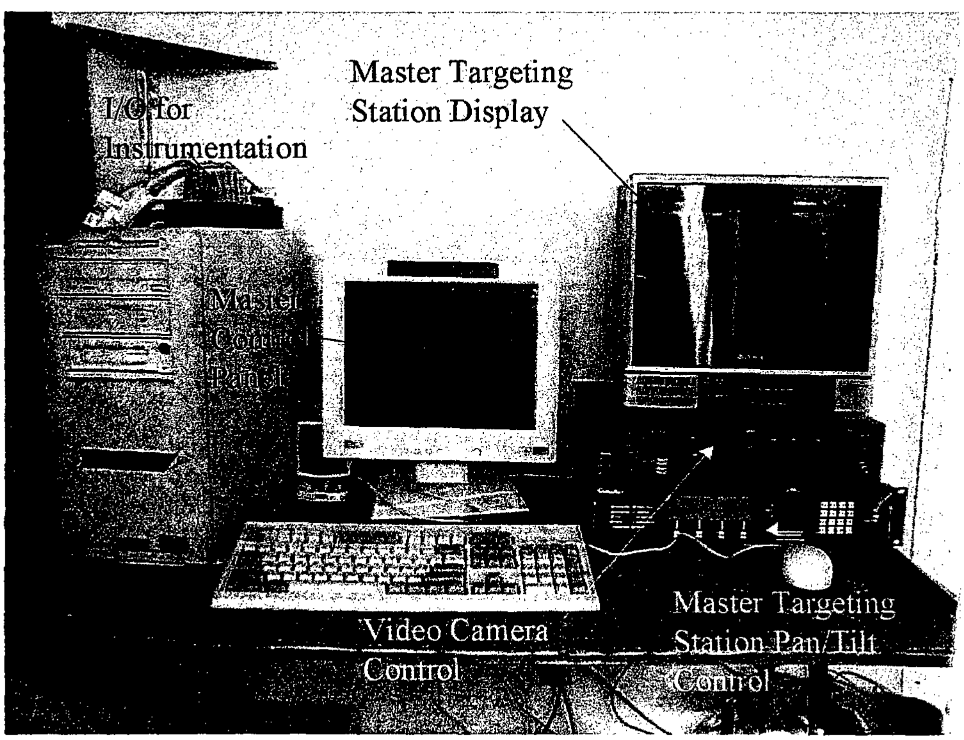


Figure 6 Control Station Layout

SYSTEM ARCHITECTURE

The instrumentation for the observer stations in the trailers communicates using an U.S. Digital SEI bus. Up to 16 SEI devices can be hooked up to a single SEI bus, each observer station has an A2 absolute encoder, an A2 absolute inclinometer, and a digital display. Using inclinometers for the elevation has advantages over using standard encoders because small differences in the azimuth planes between the observer stations does not effect the elevation measurement using inclinometers which work off of gravity. There also are SEI-M1 digital I/O modules that read the input switches in the observer stations. A diagram of a single SEI bus is shown in figure 7. Each of the observer trailers has three separate SEI busses all containing fifteen devices. The SEI busses are color-coded and all busses run from end to end of the trailers. A single power supply is needed to power each SEI bus.

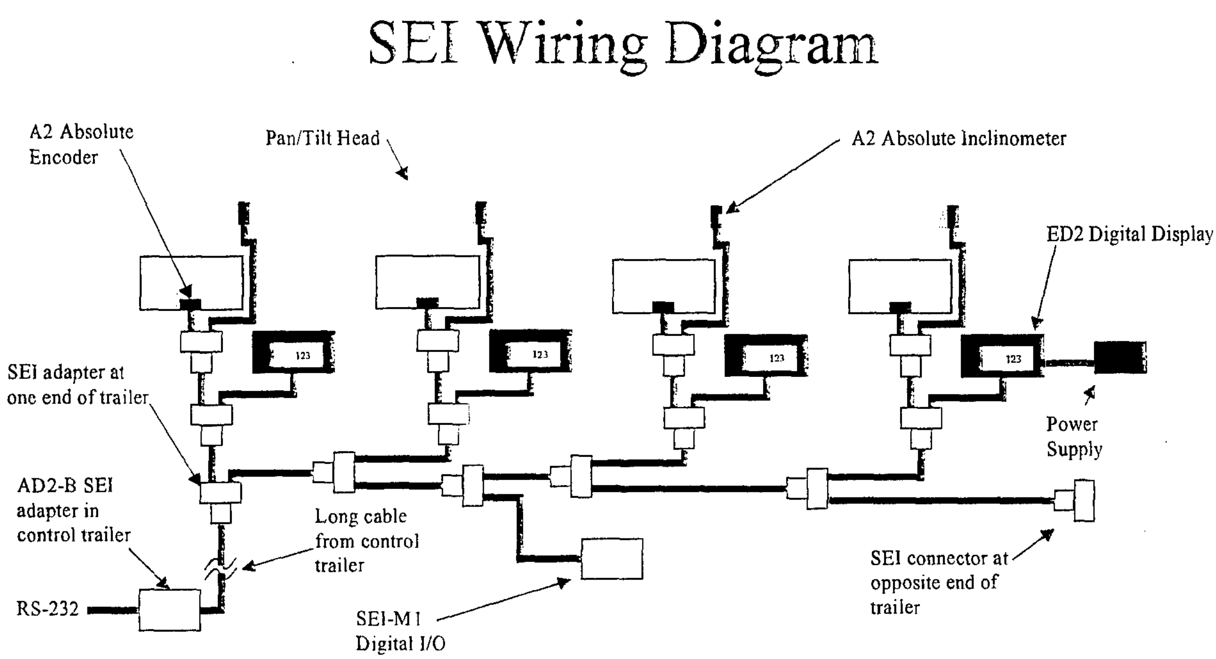


Figure 7 SEI bus diagram

Each SEI bus is connected to the control computer via an U.S. Digital AD1-B SEI adapter connected to an RS-232 serial port. A National Instrument 16-port serial breakout box is used to provide the rather large number of ports needed by the system. The input switches and indicators are wired to the SEI-M1 I/O modules and draw their power from the SEI bus through the digital display.

SOFTWARE DESCIPTION

The Observer control VI has four main sections. Initialization, Pretest Setup, Data Collection, and Verification. The Initialization section of the VI resets the value of all the variables used by the program during a run, and initializes all of the RS-232 ports used by the system. Then the program enters the Pretest set up mode which reads the current values of all of the observer instrumentation to verify that everything is functioning properly, and gives the data collector a chance to enter information about the upcoming run in the second panel of the main screen. The Target, Orientation, State, and Group # fields on the front panel come up with a selection list when clicked on with the left mouse button. This panel was set up this way so these fields would be filled out consistently to make the data easy to sort for later analysis. A screen grab with the Target selection list open is shown in figure 8. The Pretest Setup mode is also the section of the program is also were the system is calibrated. To calibrate the system, a reference target within the field of regard is selected and the master targeting station and all of the observer sights are aimed at the reference target. In the calibration panel located below the main panel of the observer control VI, there are fields to enter the distance between the trailers, reference target range. There are also buttons on the panel to calibrate the master targeting station along with all of the observer stations. The calibration aligns all

of the encoders are such that the azimuth and elevation readings are the same when all of the pan/tilt heads are aimed parallel to one another. Then the system corrects the azimuth parallax difference for each station based on the station distance from the master targeting station and the range of the current target. A screen grab of the calibration panel is shown in figure 9.

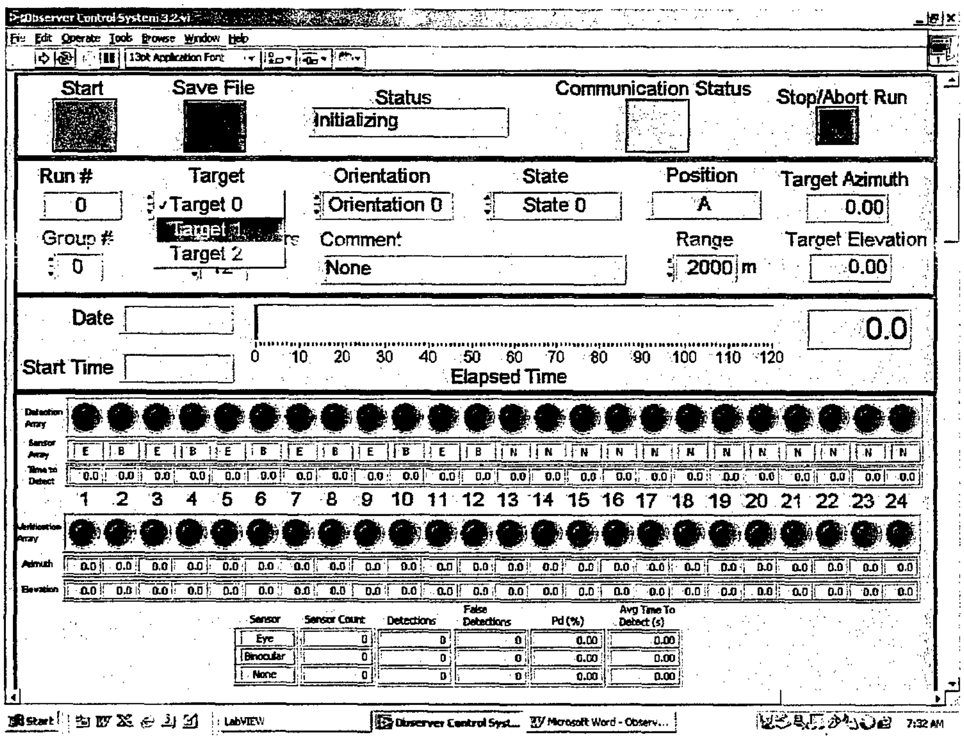


Figure 8 Screen grab of front panel with Target selection list open

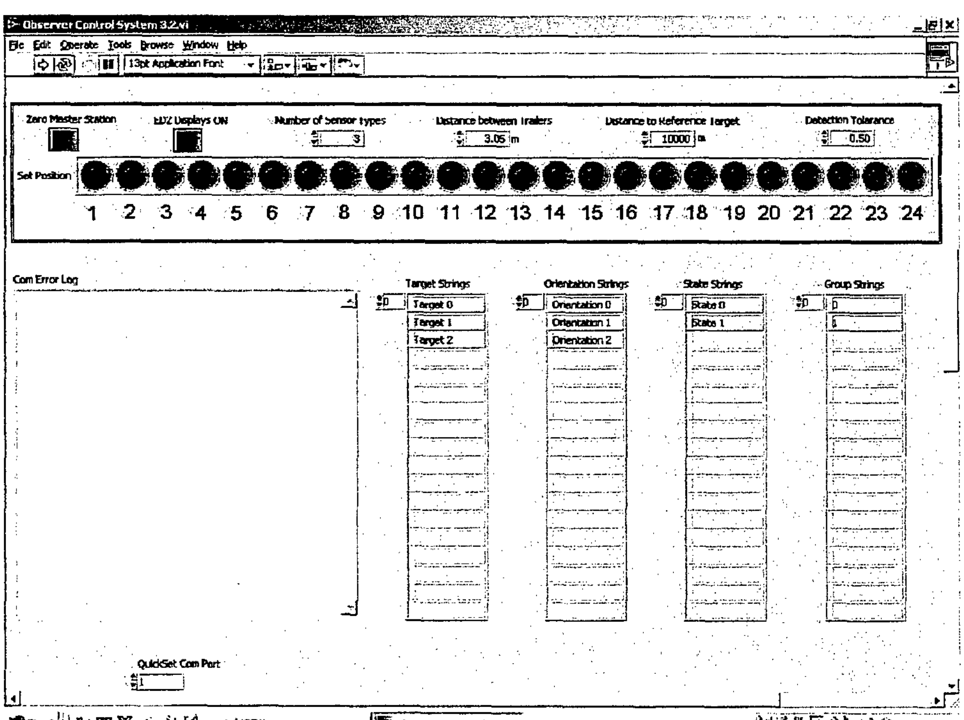


Figure 9 Screen grab of calibration panel located below main panel

Hitting the Start button in the upper panel puts the Observer Control VI into the Collecting Data mode. This will start the run timer, and log the start time and date of the run. As the test proceeds any target found and verified by the observers that is within the detection tolerance set on the lower panel will be counted as a correct detection. All of the observer actions are displayed on the main screen as the run proceeds. When the run clock has expired the program enters the verification mode, which give the observers who found a target late in the run a chance to verify their target. When all of the observers are done with their task the data from the run is saved by hitting the Save File button. Every piece of information collected on a run is saved into the data file created by the system. The data file is a tab delimited text file that can easily be imported into Microsoft Excel. There is an Excel Template with all of the columns of the data file labeled provided with the system. Every field shown on the main panel is saved in the file plus some extra information not directly shown on the screen. Most of the fields saved in the file are self-explanatory. The only new field not shown on the screen is the Detect field that shows the detection state for each of the observers. Table 1 explains what the values in the Detect field mean:

Value	Explanation
0	no detection
1	correct detection
-1	false detection
invalild	observer station empty

Table 1 Detect field explanation

The number of fields saved in the file does not change with the number of sensor types, or the number of observers. Data is saved for three types of sensors and all 24 observers no matter what the test configuration is. This eliminates the possibility of the columns of data not lining up in the file because the number of observers changed, or one type of sensor was not working for one particular run. The overall detection data from the bottom table on the main screen will be filled with zeros for any unused sensor, and the data for any unused observer stations will be filled with invalid values to make sure the data does not get included in any future analysis. Once the file is saved the VI jumps back and reinitializes and the process begins again.

THE FUTURE

Focusing on its mission of integrating Survivability Technologies, the TACOM Survivability Technology Area has given the Observer Data Collection System on long-term loan to the Chicken Little Joint Project Office. This loan makes the system available to the entire Survivability Community, although TACOM gets priority if there are any scheduling conflicts. In exchange Chicken Little is making improvements to the system. These improvements include refurbishing the trailers, which were in need of some repair, adding the capability to run tests using multiple targets in a single run, and adding GPS tracking of targets. Chicken Little is also modifying observer pan/tilts heads to accept both military FLIR and Image Intensifier sights and reconfiguring the observer stations fixing eye relief problem with rifle sight. This system is available to anyone in the DoD community that has a need for this sort of testing, and with the planned improvements the system should improve the quality and reduce the cost of doing Observer Testing in the future.

DEVELOPMENT OF A MULTI-SPECTRAL VEHICLE PROTECTION OBSCURATION SYSTEM[1]

Daniel J. Hartman, P.E., Noel Gonzalez, P.E, Edgardo Maldonado
Engineering Technology, Incorporated
Orlando, Florida
David J. Johnston
OptiMetrics, Incorporated
Bel Air, MD
William G Rouse, P.E.
U.S. Army Edgewood Chemical, Biological Command (ECBC)
Smoke and Target Defeat Team
APG, MD

ABSTRACT

Engineering Technology, Inc. (ETI) and the U.S. Army Edgewood Chemical and Biological Center (ECBC) have completed the technology development and application demonstration of a general-purpose vehicle protection obscuration system. The obscuration system employs propellant-based aerosol dissemination technology to expel obscurant materials from launched cartridges and vehicle mounted dispensers. A patent was awarded for this dissemination technology[2,3]. The smoke screen produced by this device is effective in degrading targeting sensors in the visual (VIS) and infrared (IR) spectrum. The operation of the propellant dissemination device presents no measured hazards to dismounted troops outside the visible obscurant cloud.

The smoke dispenser releases the screening aerosol directly from the dispenser unit providing concealment in the immediate area around the vehicle. The projected screening cartridges are fired from a multi-chamber launcher unit and release the screening aerosol away from the vehicle at a range of up to 300m. Each dispenser or launcher and it's electronic firing controller is integrated into a modified M2A1 ammunition container for convenience in handling, storage and use. The total weight of the prototype device is 32 lbs. The electronic controller allows the system to control start of function, the rate of smoke production, and how much of the total payload will be expelled. The controller may be set to produce a single, instantaneous cloud of very dense smoke or a continuous curtain of smoke for up to one minute's duration. This capability provides broad utility for the smoke dispenser or cartridge launcher. A fixed mounted or aimable adaptation with a combination of obscurant dispensers and launchers, in conjunction with threat warning sensors, could be employed for vehicle self-protection smoke applications.

This paper describes a proposed Vehicle Protection Obscurant System (VPOS) and the propellant dissemination technology for this application. Presentation of the paper will include video footage of live test operations and animated, three-dimensional simulations of the overall systems effectiveness.

---

[1] This work was funded by U.S. Army SBCCOM under the terms of contract number DAAD13-98-C-0043.

[2] Patent Number 6,047,644, "Propellant Based Aerosol Generation Device and Method of Use", R.J. Malecki, W.G. Rouse, D.J. Hartman, N. Gonzalez

[3] Patent Application, "A Multiple Propellant-Based Aerosol Generating Device and Method of Use", R.J. Malecki, W.G. Rouse, D.J. Hartman, N. Gonzalez

1. Vehicle Protection Obscuration System

A VPOS is a versatile countermeasure protection system providing rapid, quick response, breaklock, protection (hit avoidance), sustained protection to prevent acquisition (acquisition avoidance) and preplanned acquisition avoidance protection for maneuvering across small, short-term regions of high vulnerability. A VPOS is a required subsystem of an integrated vehicle protection system. The following remarks define a VPOS using currently available technology and, to a large degree, currently available or standard subsystems and components. The system is compatible with specialized non-lethal ammunition giving it expanded utility in military operations in urban terrain (MOUT) and crowd control situations. The VPOS and Vehicle Engine Exhaust Smoke System (VEESS), integrated with situation awareness and threat warning systems, and controlled by computerized commander decision aids, can provide manually directed, semi-automatic or automatic responses depending on the tactical situation and the commander's decisions.

The VPOS components/capabilities include the following:

- 2 Aimable, Long-Range Dischargers Protect Forward 120° Sector.
  (Provides 4 screens in each 60° Sector)
- Standard Load - 24 Long-Range RP Projectiles
  (Provides Preplanned VIS, NIR Protection)
- Vehicle Engine Exhaust Smoke System
  (Forty Minutes VIS, NIR Protection - 10 gal Fog Oil)
- 2 Aimable, Quick-Response Dischargers Protect Forward 120° Sector.
  (Provides 4 responses in each 60° Sector)
- Standard Load - 24 Time-Delay Brass Projectiles
  (Provides Rapid VIS, NIR, mid-IR (MIR), FIR Protection)
- 8 Propellant Dissemination Dispensers, 240° Sector
  (Provides Rapid or Continuous VIS, NIR, MIR, FIR)
- 3 Propellant Dissemination Dispensers, Support VEESS
  (Provides 3 Minutes Enhanced MIR, FIR Protection)

All of these components are mounted in fixed positions with manual reload. The maximum amount of obscurant material is carried in a ready-to-use condition. The majority of reloading would occur at ammunition and fuel re-supply points.

The propellant-based smoke dispenser technology described in this paper is an extension of the 40mm screening cartridge work that was previously performed by ETI.[4] The basic packaging and dissemination technology was extended to include multiple cartridges, an electric initiation method instead of a percussion primer, and a self-contained control system. ECBC's Target Defeat Team conceived a concept by which the 40mm VIS/IR/MM wavelength screening cartridges are incorporated into a stand-alone sequencing and firing mechanism to provide a continuous screening or rapid screening capability.

2. Propellant Dissemination Smoke Dispenser

The smoke dispenser housing is a standard ammunition canister approved for storage, handling and commercial shipment of this class of ordnance. A solids model of the smoke aerosol dispenser unit is

[4] Mills, T.E, "Development of a Cartridge for Aerosol Dissemination", Proceedings of the Smoke/Obscurants Symposium XX, ERDEC-CR-270, December 1998, pg 231-240

shown in figure 1. This unit is 11 inches long, 7 inches high and 5.7 inches wide with a maximum weight of 32 pounds when filled with brass powder for infrared (IR) screening. The dispenser unit contains 18

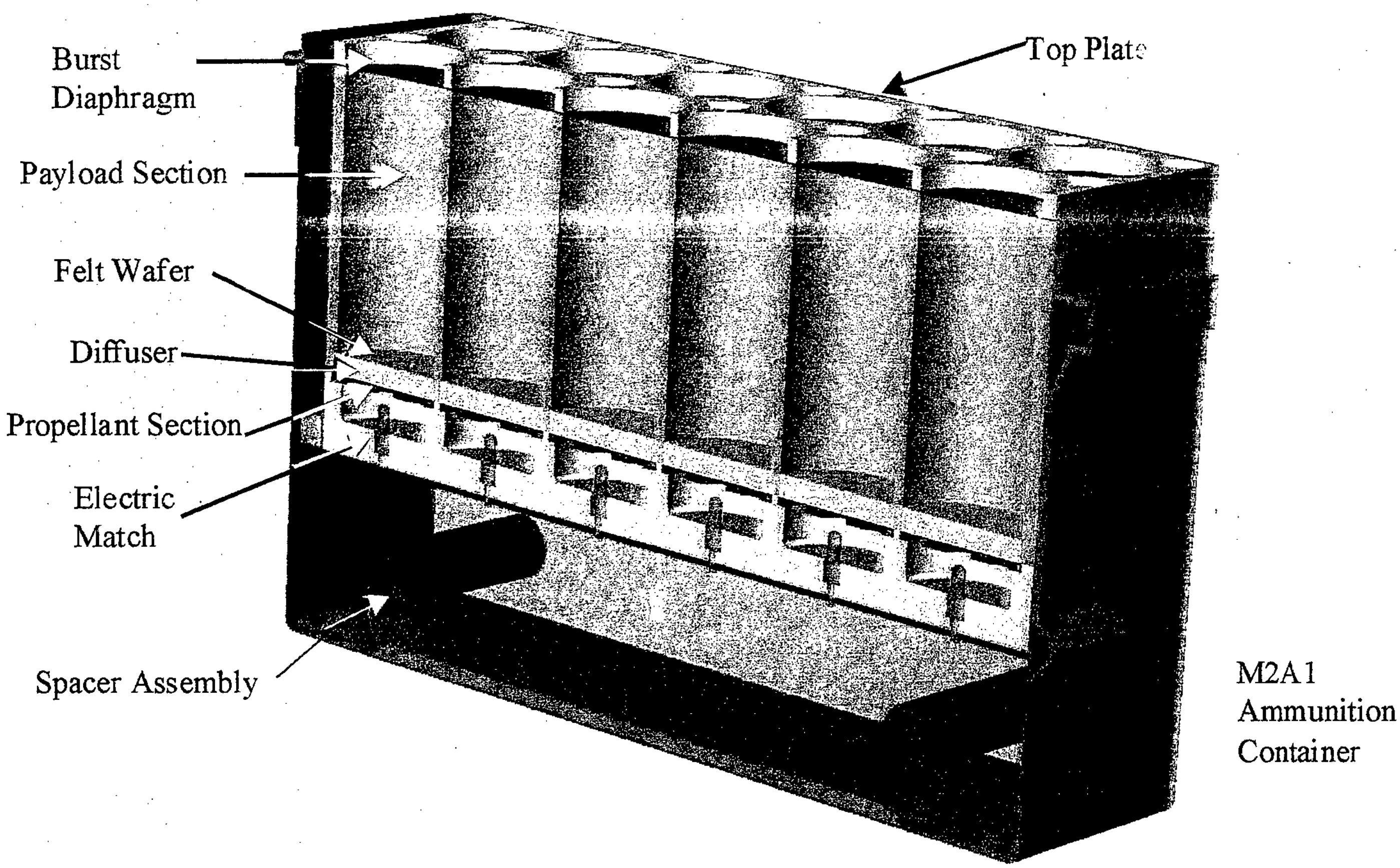


Figure 1. Smoke aerosol dispenser unit assembly configuration.

separate chambers in a vertical orientation. The construction of these chambers is generally consistent with the prior 40mm VIS/IR/MM cartridge design cited in reference 4.

At Eglin AFB a number of trials were conducted wherein the smoke dispensers were operated in a rapid pulsing fashion so as to produce a single dense composite puff to provide a rapid response obscuration function. In these trials, six chambers at a time were functioned with a firing interval of 0.1 second between each sequential chamber initiation. In this manner, three trials were obtained from a single smoke dispenser. The resulting composite puff from one such IR trial is shown in figure 2.

In other trials, the controller was set to provide a longer interval between pulses in such a manner so as to produce a continuous plume result over a longer duration of time. In these instances, a firing interval of either 1.7 seconds or 3.3 seconds was employed, giving an effective function duration "burn time" of either 30 seconds or 60 seconds for a single smoke dispenser. The screening curtain produced from a sequence of three IR chambers fired at a 3.3-second interval is shown in figure 3. As this composite cloud drifts in front of the 20-foot long panel van along the viewers' line of sight, the vehicle is completely obscured from view.

Figure 2. (U) Composite puff produced from a single salvo, 6-chamber quick burst.

Figure 3. (U) Continuous curtain produced from a 3.3-second firing interval.

3. Projected Screening Cartridge and Dischargers

A projected, propellant dissemination cartridge has been developed to provide obscuration function similar to the dispenser technology at a projected standoff distance away from the source. Its end function is identical to the smoke dispensers described in the previous sections. The cartridges are launched from a multiple-shot discharger that is packaged in a modified M2A1 ammunition container. Upon ground or surface impact of the cartridge, an obscuring cloud is produced. Single or multiple cartridges, up to 12 currently, could be functioned at ranges up to 300m.

The cartridge (figure 4) has overall dimensions of 1.8in outside diameter x 11.2in long (in-flight configuration with fins deployed), weighs 1.04lb (empty) and has a cylindrical payload section volume of 121cc (35.7mm Dia x 121.4mm L). The impact fuze uses a mass inertial mechanism, in which a weight, which houses an M209 percussion primer, slides inside a bore. Upon ground impact, the slider's inertia causes it to move forward striking a firing pin which initiates the primer. The primer then ignites the propellant charge thus starting the aerosolization process. The smoke-producing fill is encapsulated within this volume by the cartridge's aluminum casing and a threaded burst diaphragm set to rupture at 1000psi nominally. The diffuser helps in the mixing of fill particles with the gas. Upon aerosolization of the particulate fill, the payload is ejected from the base of the cartridge. The felt discs, in addition to being flame retardant, wipe the bore clean of any remaining fill that did not get expelled from the cartridge.

For flight stability, the cartridge uses a retractable fin assembly. The fins are spring-loaded to allow the cartridge to be loaded into the launch tube. Upon ejection from the launcher, the fins spring out radially and slide rearward, providing the necessary stability.

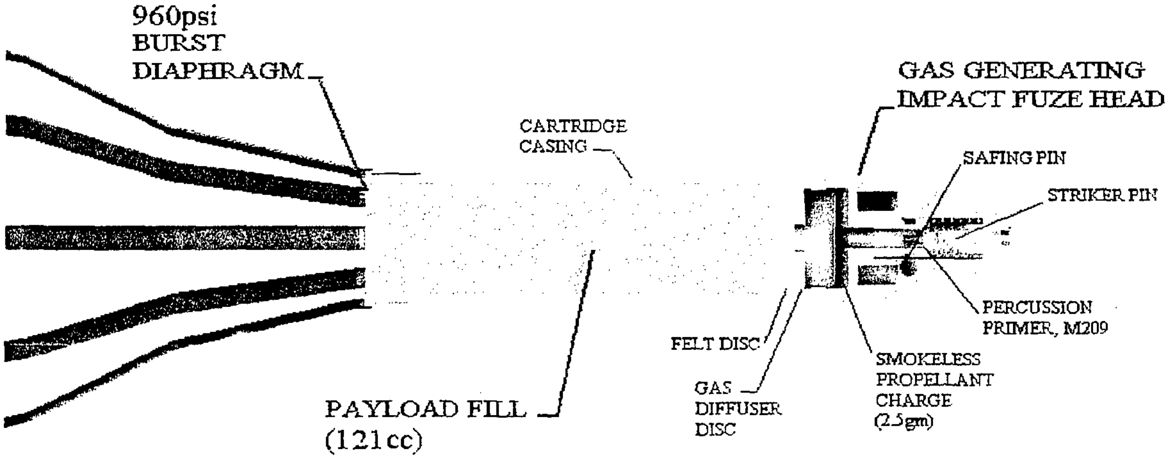


Figure 4. (U) Section view of projected cartridge.

A variation of the impact functioning projected cartridge is under development. This cartridge uses a programmable electronic delay system. The electronic delay replaces the impact fuze employed in the current design with minimal alterations to the existing hardware. The payload section of the cartridge remains unchanged. The delay is capable of functioning in the 0.1 to 9.0-second range, with an accuracy of 0.0001 sec to provide air burst functioning of the round in flight. The application of such a cartridge would be used in all range applications to provide precise positional control over cloud formations. A concept of the electronic delay system is shown in figure 5. The user can set the delay time from the control unit immediately prior to ejection from the launcher device. This allows programming individual cartridges to function at different times and also provides greater operational safety.

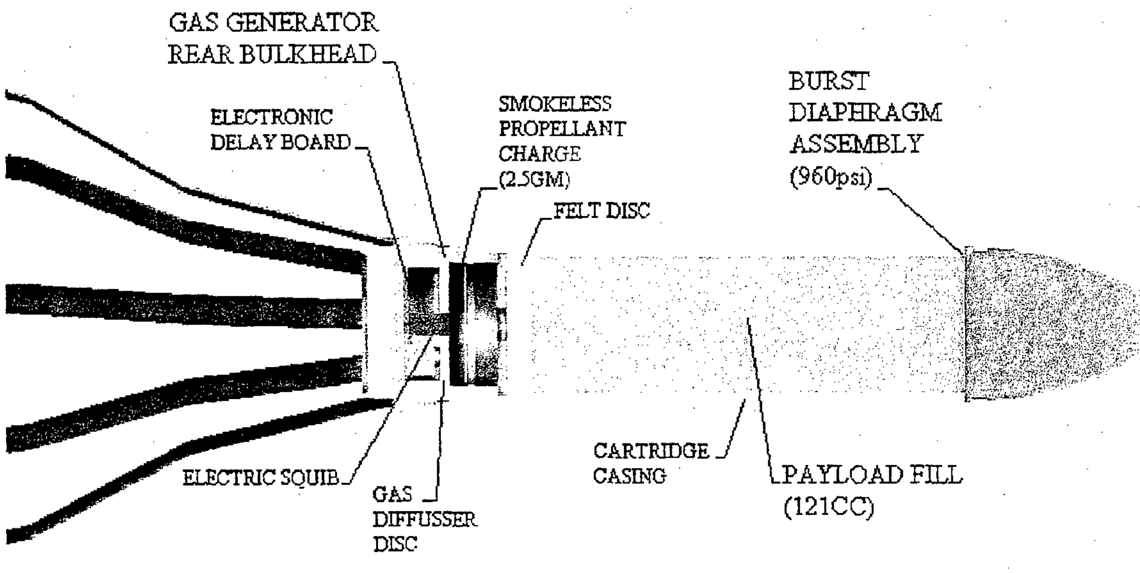
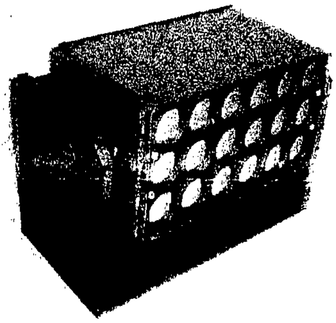


Figure 5. (U) Electronic timed delay fuze head for projected screening cartridge.

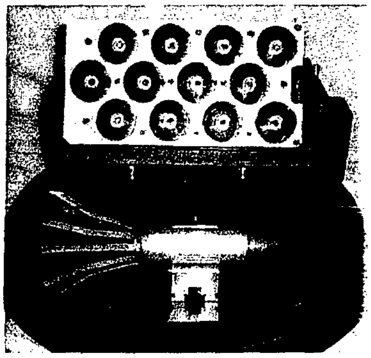
5. Vehicle Protection Implementation Concept

To provide ground vehicle self-protection against direct-fire combat, the propellant dissemination dispenser and screening cartridge discharger system could be used in an onboard, mobile pre-planned or rapid screening application. The dispenser and discharger hardware could be adapted to either a fixed installation mounting or a two-axis motor driven turret. Dispenser hardware concepts and mounting arrangements are shown in figure 6. In these applications, the firing control is integrated with the vehicle threat detection sensors and countermeasures system. Empty dispenser units may be replaced quickly in the field. It would also be possible to devise an autoloader for this application.

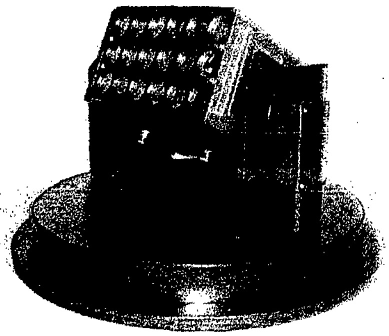
A scenario was devised to illustrate the simulation of the VPOS against visual and IR guided threats. The scenario is illustrated in figure 7. The wind is coming from the west at 3 m/s. The vehicle is traveling north and approaches an open area it knows is under surveillance by the enemy. To avoid acquisition at time "0 seconds" a salvo of three long range 40mm RP projectiles are fired to provide visual and NIR protection. At +9s the cloud begins to form and at +10s the vehicle begins its maneuver to cross the open area. At +14s the VIDS detects an IR quided missile attack to the right of the protected area so the rapid response projectiles are aimed to the left 15° and a small breaklock protective screen from 3 IR projectiles is formed at +15s. As the vehicle continues to maneuver to cover, IR reacquisition prevention obscurants are functioned at time +18s, +23s, and +28s, and a 1-minute sustained IR screen is initiated at time +34s. The vehicle continues to maneuver to counterattack and continue with its original mission.



Fixed Installation Mount



Projected screening cartridge discharger unit.



Two Axis Motor Driven Turret

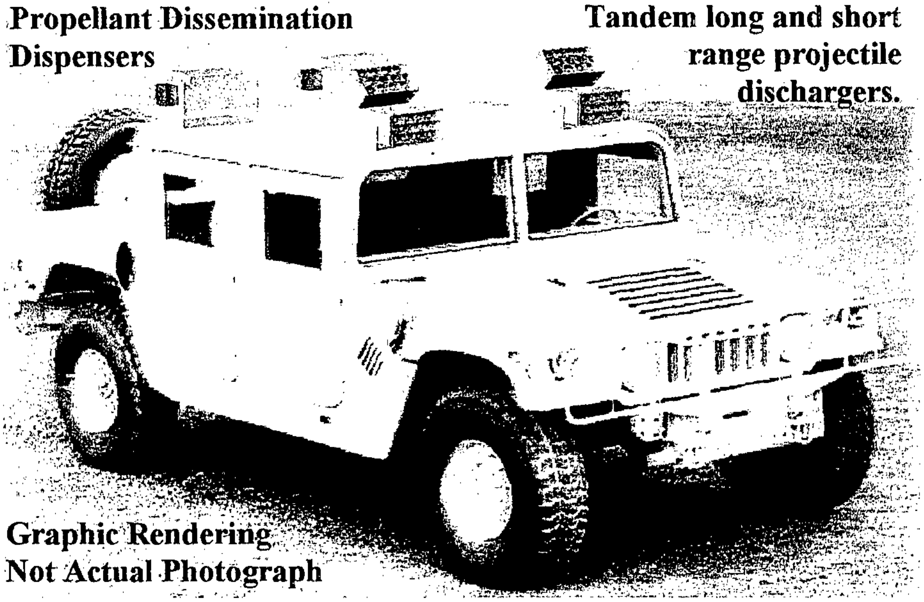


Figure 6. (U) Vehicle self-protection obscurant countermeasure application.

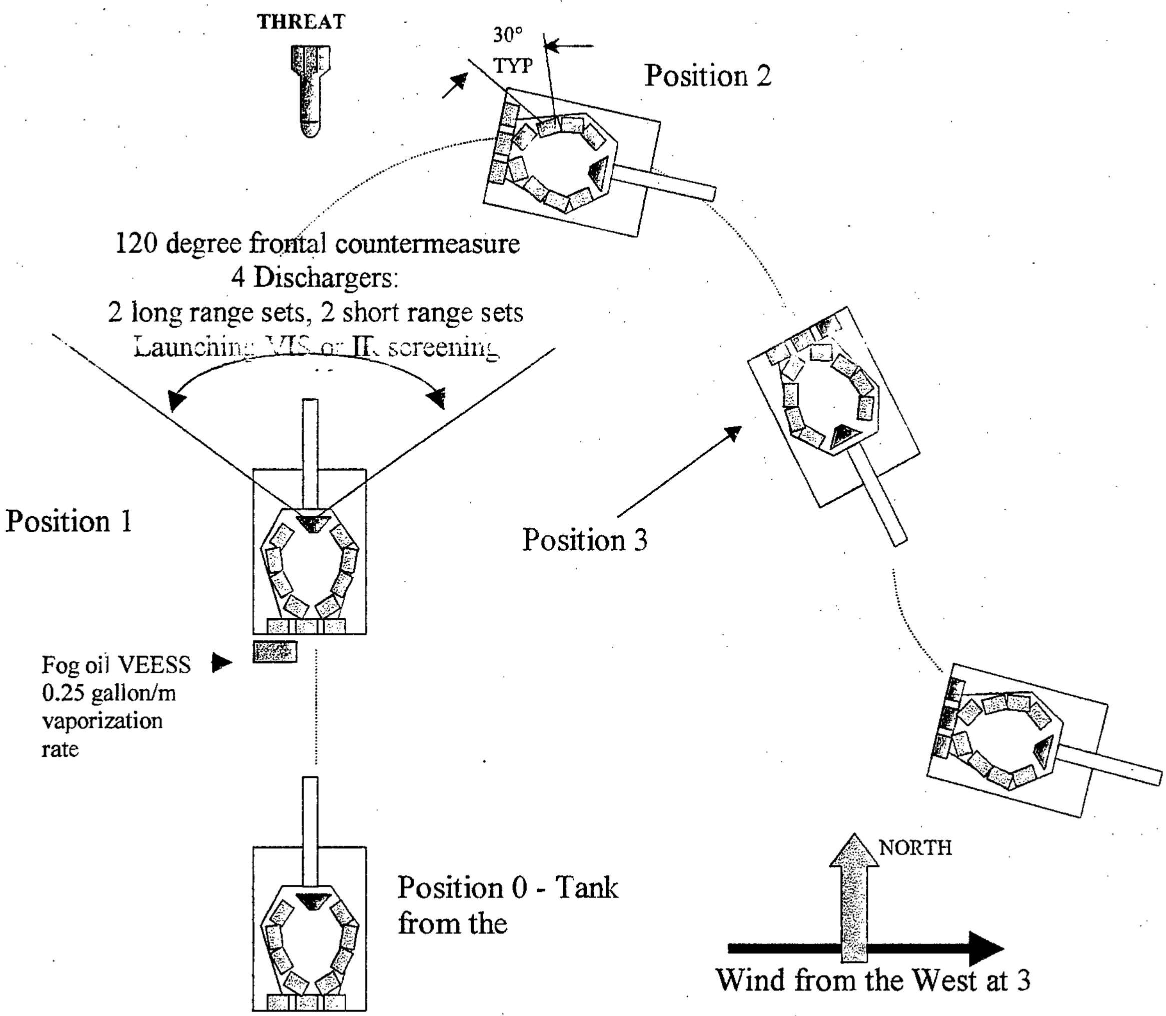


Figure 7. (U) Diagram of vehicle protection system simulation Scenario 1.

Selected frames from the dynamic simulation results are provided in figure 8. In these images, the visible white smoke produced by the RP screening cartridges is shown in white. The IR screening aerosol produced by the brass flake fill in the cartridges and the vehicle protection dispensers are represented as a brown aerosol. In figure 9 the Cloud Density Visualization Utility (CDVis) for SSPM represents the cloud images as seen when viewed in a selected spectrum; VIS, NIR, MIR or FIR. The effectiveness of the visual smoke decreases across the IR spectrum. The screening effectiveness of the brass flake aerosol does not change across the VIS to FIR spectrum because the extinction coefficient for brass flake is generally constant across this spread of wavelengths. The extinction coefficient is an optical property of material particles which is representative of the aerosol screening cross-section per unit mass at a specified wavelength. As illustrated in figure 9, the target vehicles own thermal sights would not be significantly impaired by the visual screening smokes. However, the IR smokes are opaque across the entire visual and IR spectrum. In the simulation, the long-range projectiles establish an initial VIS/NIR acquisition prevention screen. The IR capability of the enemy allows attack with a more sophisticated weapon system, but the VIDS warning systems directs the rapid burst of projected screening cartridges and quickly establishes a curtain of IR protection against the incoming threat. The initial dispenser obscurant releases from the left-side dispensers provide reacquisition screening protection directly coupled to the vehicle followed by sustained IR protection from the continuous dispenser function as the target vehicle moves clear of the vulnerable area and proceeds with counterattacks.

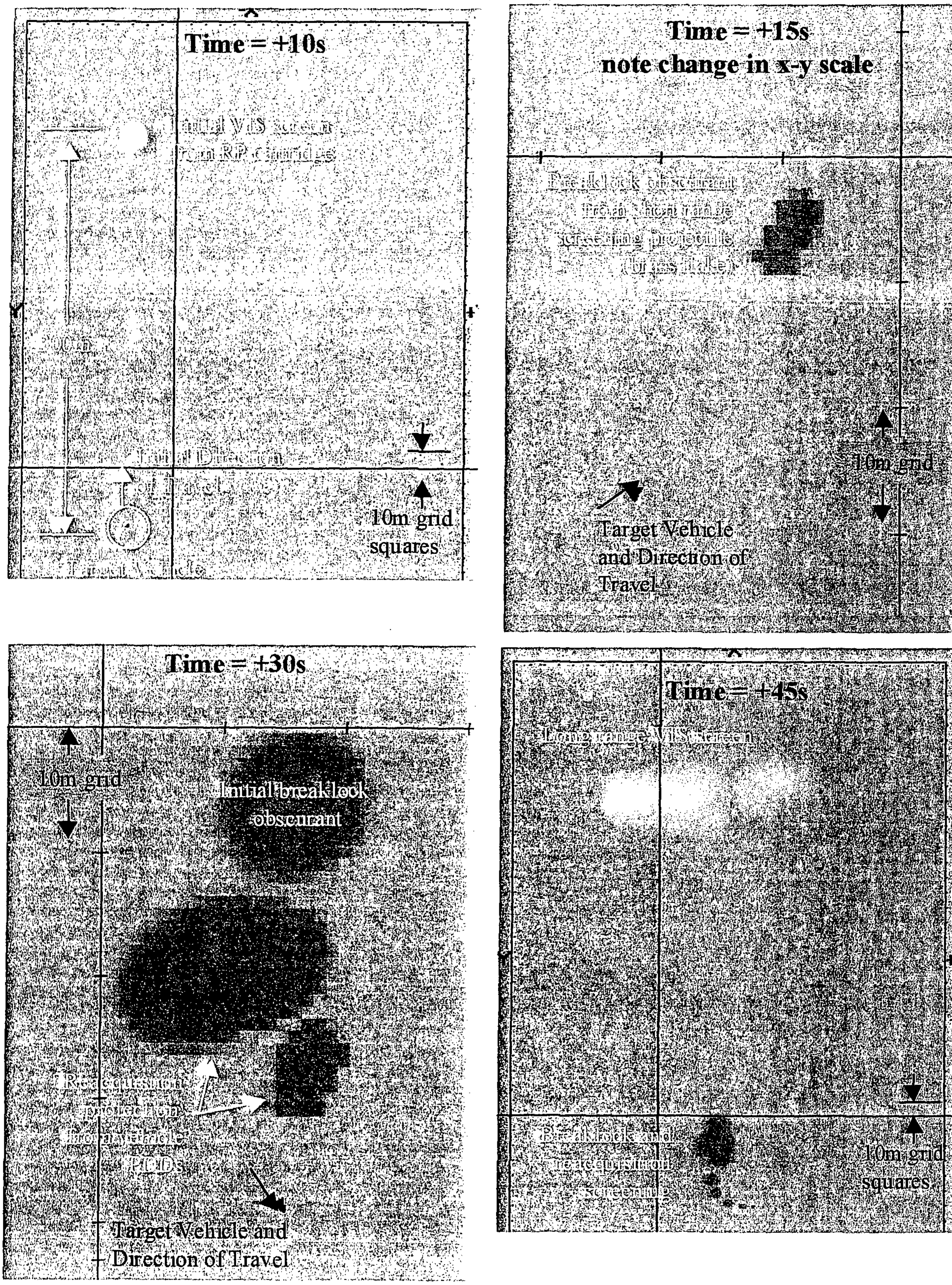


Figure 8. (U) FCS vehicle protection visual screening scenario simulation results. Time sequence of screening aerosol release from SSPM simulation.

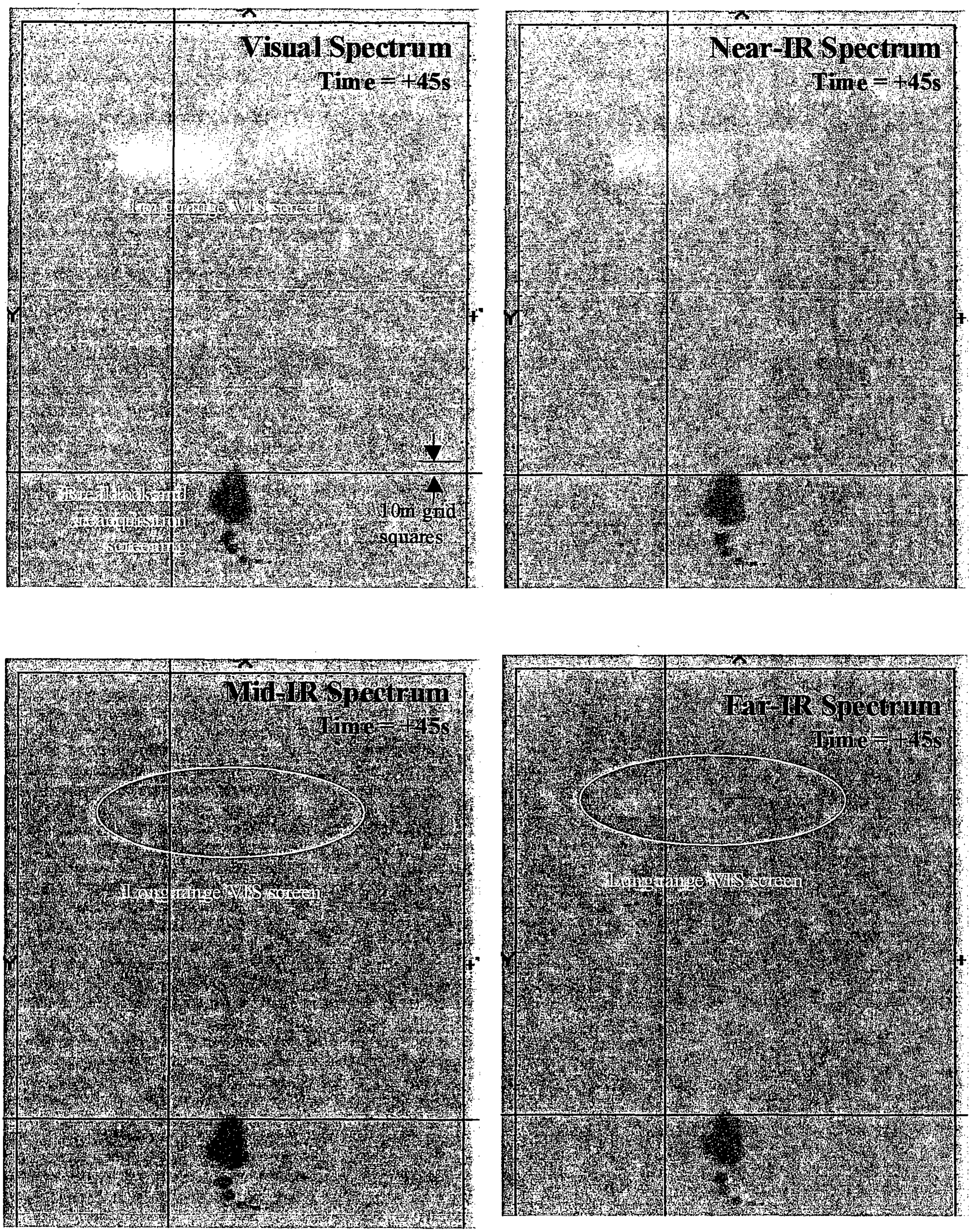


Figure 9. Visual and IR smoke as seen in different spectrum.
Simulation time T0 + 45 seconds.

6. Conclusions

The emphasis for the Interim Army Vehicle (IAV) and Future Combat Systems (FCS) is on rapid deployment, lightweight combat vehicles. Propellant dissemination technology for obscurant countermeasures affords a significant opportunity to replace heavy armor with a low-cost, versatile, multi-spectral and effective vehicle protection suite with the capability to deploy a rapid sensor guided missile countermeasure to achieve break-lock. A comprehensive Electro-optic-Infrared (EO-IR) countermeasure suite will include active electronic detectors and jammers, decoys and obscurants. The propellant dissemination methods and low hazard dispenser technology is dually suited to dispense non-lethal aerosols in urban terrain. The propellant dissemination technology described in this paper is available now.

Passive Infrared Signature Augmentation of Full-Scale Plastic Targets

Lisa M. Gebus, Jeffrey S. Sanders
Signature Research, Inc.
150 West Park Loop, Suite 203
Huntsville, AL 35806

ABSTRACT

The Targets Management Office (TMO) manages the development, acquisition, and operation of both aerial and ground target systems for use in destructive and non-destructive testing by the U.S. Army T&E community. A need has been identified for low-cost, full-scale validated targets that can accurately simulate the visual, infrared (IR), and radar signatures of threat systems. To address this need, a program was initiated by TMO to augment an existing full-scale, vacuum-formed plastic target with sufficient signature fidelity to adequately stress U.S. Weapon System sensors. This paper discusses the portion of the program dealing with the development of a passive IR system to meet signature needs. The concept presented in this paper will detail a cost effective design solution that augments the plastic targets with passive thermal properties that emulate actual vehicle armor throughout the diurnal cycle.

The chosen solution for passive IR signature emulation was the addition of water jackets to these plastic targets to replicate the thermal mass of actual armor. This technology not only is an affordable approach to improving the passive signature of the existing plastic target, but also is easily supportable on a test site. The water jacket concept has previously been proven in the BMP-3 Surrogate program. In this program, the amount of water necessary to imitate the thermal mass of the actual target was calculated. Trade studies were then conducted between signature fidelity and structural integrity to arrive at a solution that would provide sufficient IR signature while still allowing for the plastic target to support the additional weight of the water. The combined weight of the plastic shell and water jackets is significantly less than that of the actual target, but the plastic target has an effectively equivalent thermal mass of a heavily armored vehicle.

The process for implementing water jacket technology on plastic range targets is presented. Modeling and simulation techniques were prevalent in this program. A meshed model of the plastic target was created and material properties and thicknesses of both the shell and water jackets were incorporated into this model. Thermal simulations were then conducted to compare the thermal properties of the plastic target to actual vehicle thermal properties in order to achieve the optimal placement and configuration of the water jackets on the plastic target. A sub-structure support system was then designed to withstand the load of the water jackets. The water jackets were then fabricated and attached to the plastic shell, and field-tests were conducted to verify the signature fidelity and structural integrity of the surrogate target. IR signature comparisons of the surrogate against the actual threat target are presented to demonstrate the feasibility of this approach.

INTRODUCTION

The Targets Management Office (TMO), Project Manager for Instrumentation, Targets, and Threat Simulators (PM-ITTS), U.S. Army Simulation, Training and Instrumentation Command (STRICOM), is responsible for the development of ground targets for weapon system testing. The requirements for the targets are customer driven and many times these requests are for full-scale, low cost targets that accurately represent the visual, infrared and radar signatures of threat systems. In order to accurately stress the weapon systems, it is crucial to have a ground target that is a correct representation of the intended threat system.

The Naval Air Warfare Center Atlantic Targets and Marine Operations of Patuxent River Naval Base has developed a low-cost vacuum forming process to fabricate a variety of plastic armored vehicle targets. These surrogate targets support a wide range of testing and training requirements. The targets are created out of 5/32" thick ABS plastic sheets and the plastic is available in either olive or desert sand or can be painted in camouflage colors. The vehicles are usually mounted on skids and can easily be towed behind a pickup truck. Visually, these targets are an accurate representation of the actual target for which they are modeled after. Figure 1 shows a photograph and a thermal image of the T-72 plastic target.

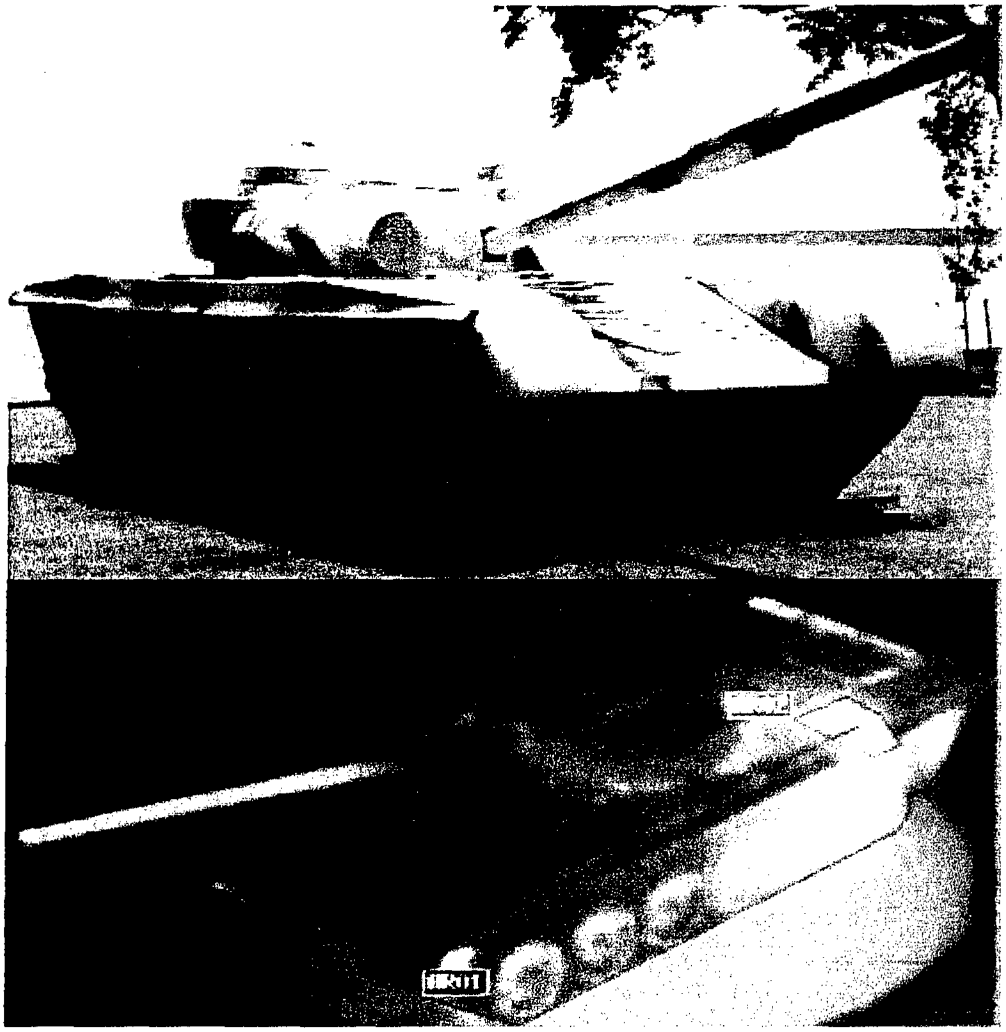


Figure 1. Vacuum-formed plastic T-72 photograph and thermal image.

The plastic ground targets meet the requirements of being a low-cost, visually representative target. In addition to being visually representative, TMO customers require a degree of IR and radar signature accuracy to effectively stress their sensor systems. This paper will concentrate solely on emulating the passive IR signature. The plastic targets created at the Patuxent River Naval Base need to be augmented in order to correctly simulate the passive IR signature of the actual vehicle. The augmentation process consists of the addition of water jackets to the underside of the plastic targets to replicate the thermal mass of thick armor present on an actual threat system.

Currently the plastic targets can be augmented to provide an active IR signature of the engines running and the barrel heated from recent gun flash. The thermal image in Figure 1 is of the plastic target with the active signature augmentation hardware in operation. This augmentation is the result of heating elements added to areas of the target such as the exhaust port, wheels, and gun barrel. The heating elements are used to represent the target in specific exercised conditions, but aren't capable of accurately portraying the vehicle when outside factors such as, time of day, wind and solar loading become of interest. The existing augmentation process neglects representation of the thermal mass of the target, which is important in simulating a correct IR signature. The water jacket augmentation process will generate a passive IR signature that is responsive to the effects of the surrounding environment, much the same as an actual vehicle would respond under natural conditions. The threat system of interest for this program is the T-72 main battle tank. The thick armor of the T-72 has a large thermal mass. The temperatures of objects with large thermal masses will respond more slowly than plastic to natural heat inputs such as

wind and solar loading. Therefore, the thermal mass of a target must be replicated in order to match surface temperatures over the course of a day and in different environmental states.

The process for implementing water jacket technology on plastic range targets is shown in Figure 2. This process includes creating the vehicle geometry, creating a meshed model, attributing it with material properties, calculating and adding water jacket thicknesses, and running a thermal simulation to compare the augmented plastic target model to an actual vehicle model. After the thermal simulations are completed and the water jacket thickness is determined, a sub-structure support system is then designed to withstand the calculated weight of the water jackets. Each water jacket is then fabricated and attached to the pre-existing plastic shell. Field-tests are then conducted to verify the signature fidelity of the surrogate against the actual target.

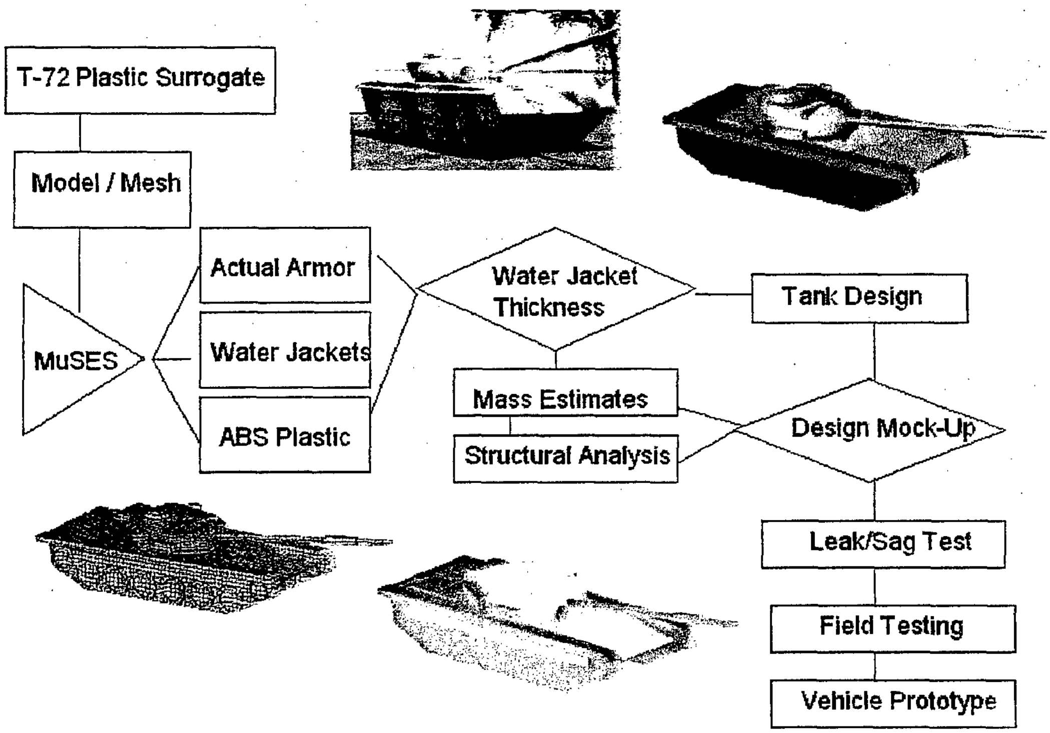


Figure 2. Plastic target augmentation process.

AUGMENTATION PROCESS

The entire process for designing water jackets begins with determining which sections of the plastic target require a water jacket to accurately simulate the thermal mass of the actual target. Thicker sections retain heat longer and therefore require thicker water jackets to simulate their thermal mass. Thinner sections, such as the gun barrel, fender skirts and turret boxes, do not require water jackets. After analysis of the T-72, the following sections require water jackets to simulate their thermal mass: glacis, hull sides, hull top, hull rear, turret top, turret front and sides, and turret rear. These locations are shown in Figure 3.

The critical parameter for the augmentation process is the thickness of the water jacket. Since the length and width of the water jacket and armor section have to remain the same, the thickness needs to be calculated to yield an equivalent heat capacity for the appropriate volume of the actual material. The derivation of the equations used to calculate the water jacket thickness has been published previously. [1]

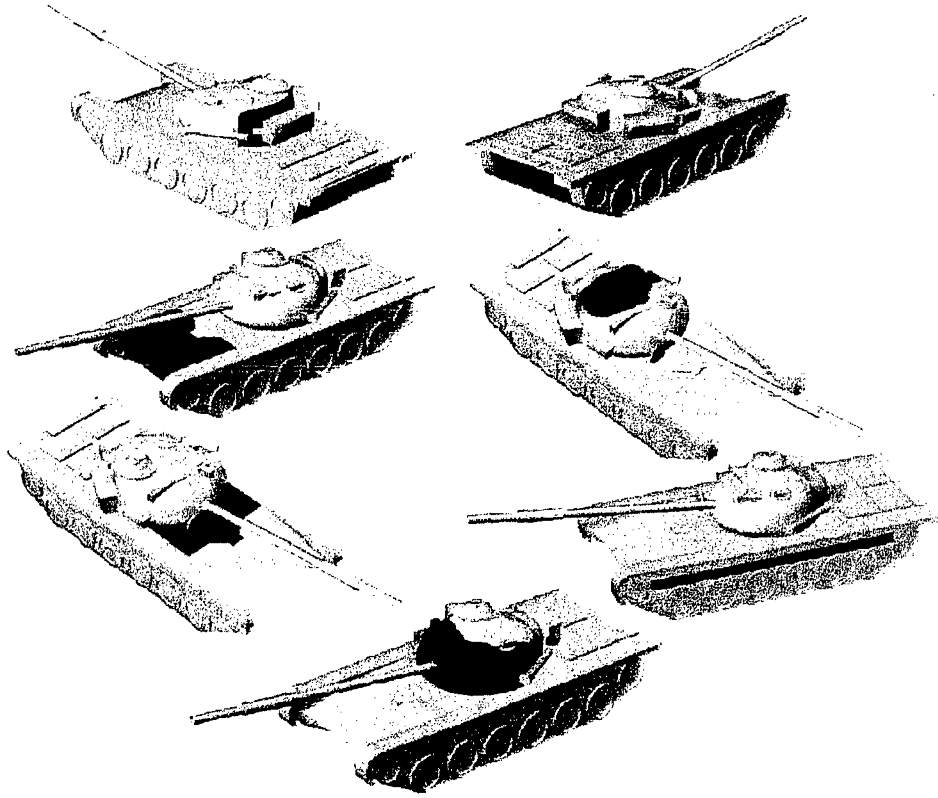


Figure 3. Water jacket locations.

The goal of the calculation is to determine what amount of water is necessary to emulate the thermal mass of actual armor. Given the composition of the actual armor, the density and specific heat of the material(s) must be known in order to complete the calculations. The water jacket wall thickness is governed by the thickness of the material used to build the tank. The water jackets of the T-72 are vacuum formed from the same material as the pre-existing plastic surrogate shell. This material is 5/32" ABS plastic. For the purposes of calculation, the imperfect thermal conduction between the ABS and water is ignored and the heat capacity of the ABS and water is calculated as a thermally homogeneous mass. The thickness of the water is the only unknown. The water thickness can be calculated by the following equation:

$$d_{water} = \frac{d_{armor} \cdot \rho_{armor} \cdot C_{parmor} - 2 \cdot d_{ABS} \cdot \rho_{ABS} \cdot C_{pABS}}{\rho_{water} \cdot C_{pwater}} \tag{1}$$

where d_{water} is the water thickness, d_{armor} is the armor thickness, ρ_{armor} is the density of the armor, Cp_{armor} is the specific heat of the armor, d_{ABS} is the thickness of the ABS, Cp_{ABS} is the specific heat of the ABS, ρ_{water} is the density of the water, and Cp_{water} is the specific heat of the water.

The next step in developing water jackets for the plastic surrogate is to create a computer model of the target. Dimensions for the model were taken from an existing plastic surrogate located at Redstone Arsenal. The turret was based on a full-scale actual T-72 model that was previously created. The model generated is a simplified version of the plastic surrogate. Details, such as the grating on the radiator and additions to the glacis, were left off. By decreasing the amount of unnecessary detail, the time to complete a thermal simulation decreases.

Once the model is completed, it is properly meshed to allow for conductive heat transfer between thermal nodes. The meshed model is imported into the Multi-Service Electro-optic Signature (MuSES) code and each part is attributed with the appropriate material properties and thicknesses. The vehicle model is attributed as the actual target, as a plastic shell, and with water jackets to determine the improvement the water jackets make over the pre-existing plastic surrogate and to

determine how close the water jacketed target is to the actual target. Several simulations are run in MuSES under different weather conditions to indicate how the vehicle reacts under various solar loads. For this reason, both cloudy day and sunny day weather files are run in simulations. The results from the simulations are compared for each part and the improvements are quantified. After the results of the simulation are evaluated, the optimal thickness for the water jacket can be determined. Trade studies are conducted between signature fidelity and structural integrity to arrive at a solution that provides sufficient IR signature fidelity while still allowing for the plastic target to support the additional weight of the water.

With the thickness of the water jacket finalized, the additional mass of the water can be calculated. This is accomplished by knowing the density of water and the volume of each water jacket. The mass of the water jacket assembly is critical in designing the supporting framework. The supporting structure is designed using common lumber and fasteners. This keeps the materials both inexpensive and easily accessible. The design of the structure is made to be as simple as possible for ease of assembly, yet strong enough to safely support the load from the water.

Once all the design work is completed, the plans can be given to the fabrication shop so the parts can be fabricated. The first step to fabricating the water jacket is to build a mold in the shape of the jacket. The mold is built out of layers of dense particleboard. Each mold has to be extremely strong to withstand the pressure from the vacuum forming process. Large sheets of ABS plastic are vacuum formed around the molds to create the shape of the water jacket. The particleboard is porous enough to allow the vacuum to pull air through it. Once the mold is formed from the ABS sheet, the excess material is cut away, leaving the desired water jacket. The exterior plastic parts are also created in this manner. Examples of the molds can be seen in Figure 4.

Exterior Turret Molds

Turret Water Jacket Mold

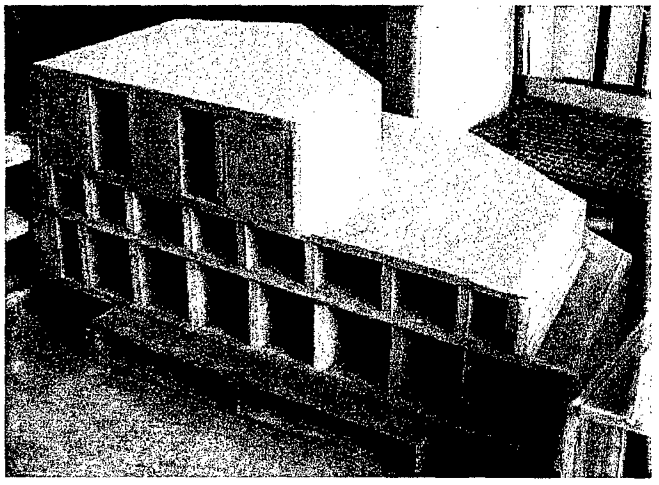


Figure 4. Vacuum forming process molds.

The water jacket molds are glued to the exterior molds using a mixture of methyl ethyl ketone (MEK) and ground up ABS plastic. For additional sealing a plastic welding process is used. A special welding gun melts the two surfaces to be welded and injects a stream of hot ABS into the joint, bonding the two surfaces. For the thicker water jackets, internal support is necessary between the water jacket and external skin. Several different methods have been tested. I-beams created from vacuum formed 2x4's are used inside the turret. Another internal support structural concept is using plastic rods drilled through both layers.

The glacis water jacket was the first thick water jacket to be built. The initial fill test showed evidence of leaks in various areas. These leaks were addressed with additional MEK solution and layers of fiberglass cloth adhered to the surface with MEK solution. The second fill test proved to be successful and the water jacket successfully held water. The surface

temperature of the glacis was noticeably cooler to the touch than the fenders under direct solar loading at approximately 10:00. This initially demonstrates that the theory of water jackets is successful. The glacis, complete with water jacket, is shown in Figure 5.

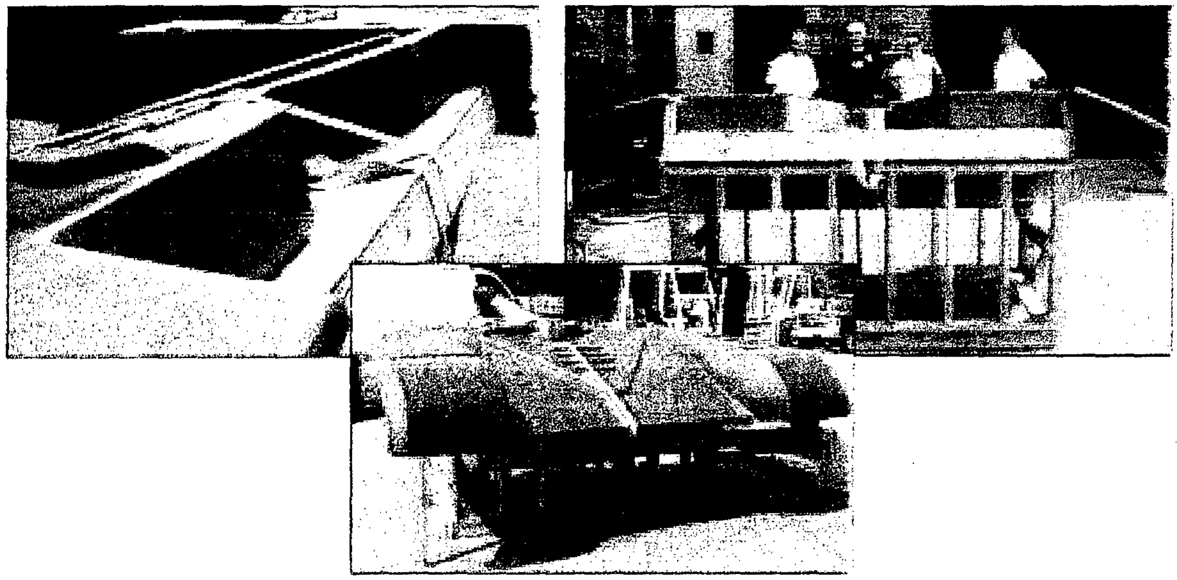


Figure 5. Glacis water jacket testing.

Field test plans are developed to verify the results of the analyses and simulations, and to ensure structural integrity of the design prior to final prototype fabrication. The purpose of the field test is to measure and quantify any thermal differences between the actual and augmented plastic target. The goal is for the plastic surrogate to display passive thermal properties that emulate actual armor systems throughout the diurnal cycle. Thermocouple data will be sufficient for comparison between the actual, surrogate, and augmented surrogate. The instrumented vehicles will have data recorded as temperature vs. time. All vehicles will have to be measured simultaneously and the environmental histories of the vehicles will need to be identical. All test vehicles need to be at the test site a day in advance of the first measurements so that they reach thermal equilibrium with the local environment. The water jackets of the plastic target should be filled at least 24 hours prior to data collection to allow the water temperature to come to equilibrium with the surrounding environment. Initially, temperature data is used for comparisons of IR signature for the surrogate against the actual threat target. At first, only the glacis of each vehicle will be instrumented and tested to initially validate the process. Field testing of the entire vehicle will follow, with the addition of infrared image measurements for full validation.

SAMPLE RESULTS

Figure 6 displays the estimated mass calculations. The water jackets, when full, will add approximately an additional 4000 pounds to the current plastic surrogate. For transportation purposes, the water jackets will be drained and the mass of the target will not include the additional 4000 pounds. However, weight will also be gained from the additional wooden structure necessary for supporting the load of the water.

The combined output of the diurnal cycle simulations for three vehicles is displayed in Figure 7. An identical node on the glacis was selected for each of the three simulations. The temperature of that node throughout the diurnal cycle is recorded and graphed. The simulation begins prior to sunrise, and continues until that time the following day. The weather file used in obtaining these results was a sunny day on July 19, 1984, with latitude of 47.175 and longitude of 88.492. The approximate sunrise and sunset times are 6:17 and 21:43, respectively. The vehicle is positioned facing south. The dip at approximately time equals 8 hours is due to the shadow of the gun as it passes over the turret. This is very apparent on the plastic target due to the thin ABS material. This large temperature fluctuation is proof that the plain plastic target is insufficient in representing the thermal signature of an actual T-72. The effect created from the shadow shows how quickly the thin plastic cools off without the direct solar loading from the sun. The same is true for the heating effect caused by the

impact of the solar radiation. These results show the selected node of the actual target and the water jacket target track fairly closely. The addition of a water jacket to the glacis creates about a 75% solution to achieving the correct thermal signature of an actual T-72. Similar graphs have been developed for each water jacket, all displaying an improvement in the signature from the plastic to the actual target.

Estimated Mass Calculations

Part Name	Armor Material	Armor Thick mm	Water Thick mm	Water Thick in	Area in^2	Volume in^3	Mass lbs
Glacis	Laminate	200	200	8	5091	40728	1434
Hull Sides	Steel	60	52	2	2727	5454	192
Hull Top	Steel	30	24	1	2695	2695	95
Hull Rear	Steel	60	52	2	3024	6048	213
Turret Top	Steel	45	38	1.5	3522	5283	186
Turret Rear	Steel	120	108	4.25	2434	10345	364
Turret Front /	Laminate	250	229	9			
Turret Sides (2)	Steel	120	108	4.25			
			avg.	6.63	5154	34145	1202
Turret Top Sides	Steel	45	38	1.5	4964	7446	262

Figure 6. Estimated mass calculations.

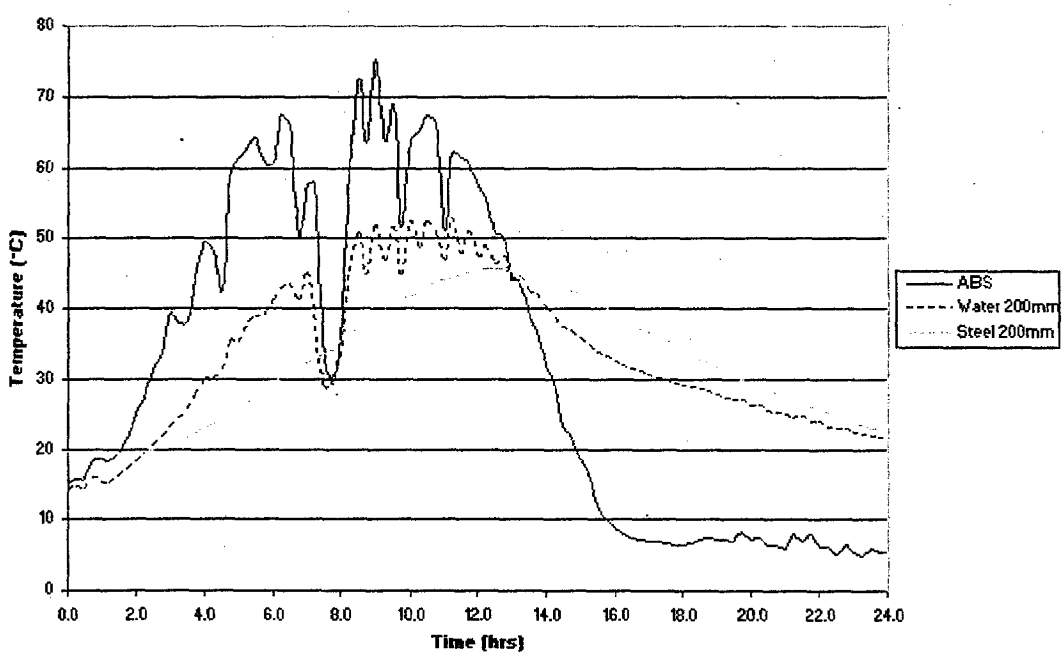


Figure 7. Glacis simulation results.

Two sets of sample results are displayed in Figure 8. The first row shows the targets at 10:25. This is approximately 4 hours after the sun initially hit the targets. As seen in the figure, the thicker areas of the glacis and turret of the real target are cool in contrast with the surrounding thinner parts. The plastic target is one homogeneous mass, lacking the internal thermal contrasts of the real target. The results for the target with water jackets are similar to those of the actual target. The water jackets delay the temperature increase from the morning sun for the thicker parts and are successful in mimicking the thermal imagery of the actual target.

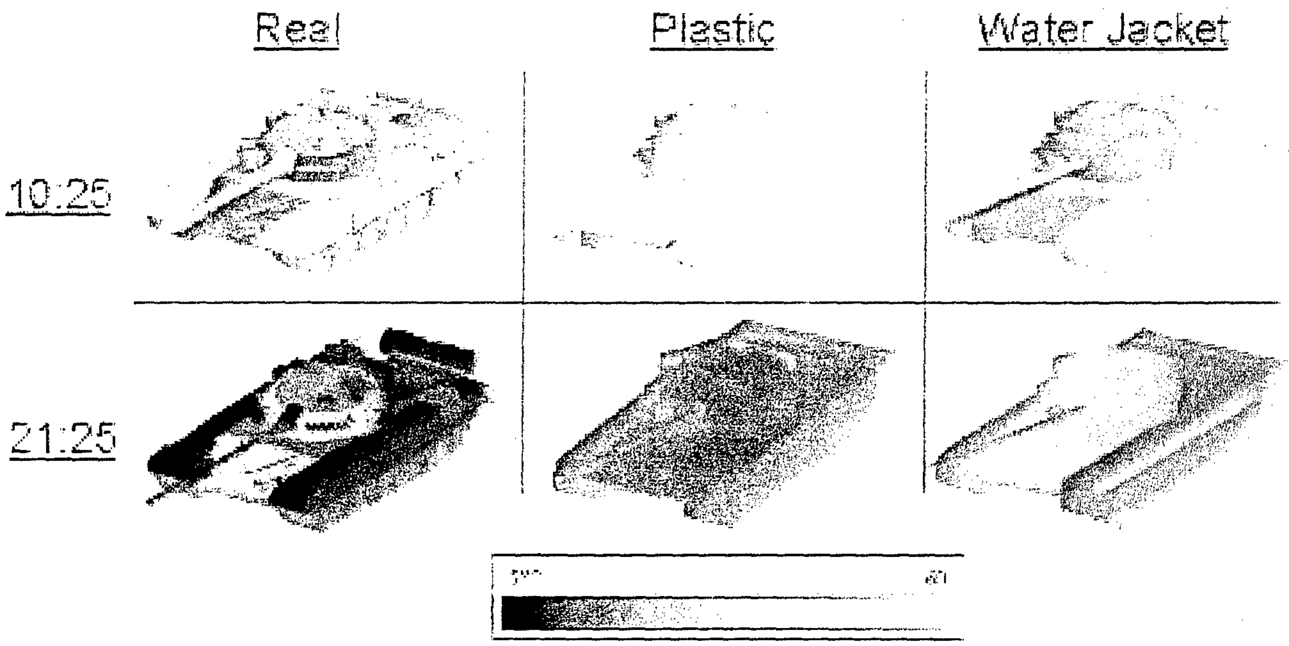


Figure 8. Sample thermal image results.

The second row shows the target at 21:25, right around sunset. The thinner sections of the real target have cooled down, but the thicker sections are still relatively warm. The same holds true for the target with water jackets. Once again, the plastic target is a homogenous mass that quickly cooled as the heat of the day decreased. These results demonstrate the utility of water jackets for plastic target IR signature augmentation.

The existing plastic target uses a wooden structure for the supporting framework. This is a fairly simplistic structure based on the fact that it only has to support the weight of the thin ABS plastic shell from which the target is manufactured. In order to support the additional weight incurred by the water jackets, it is necessary to design further support structures to withstand the load of each water jacket.

To achieve this goal, the initial framework is modified to support the larger load. The left side of Figure 9 shows the original support structure necessary to support the existing plastic T-72 ground target. Since there is a chance that people may be underneath the support structure, the structure is overbuilt to have a factor of safety of at least 2. Design selection is based on ease of assembly, supporting ability, and design simplicity. The middle of Figure 9 shows steps in the process of improving the framework. The right side of Figure 9 shows the initial supporting framework with some of the support structure additions. Each water jacket has a support structure designed to support the load of that jacket. The individual support is then integrated into the existing framework.

The augmented plastic target is intended for stationary usage only. In the event that the target has to be relocated, each of the water jackets will need to be drained. The momentum created with moving large masses of water could cause the rigid water jackets to crack and break. The surrogate target can be relocated on skids with empty water jackets much the same as the existing plastic targets.

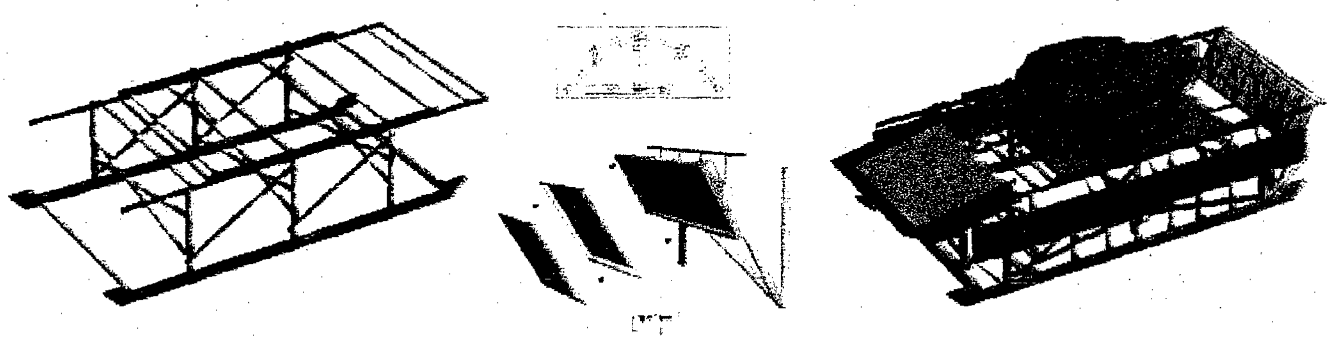


Figure 9. Additional supporting framework.

CONCLUSION

This paper presented the concept of using water enhanced plastic targets to satisfy TMO's needs for low-cost, full-scale, ground targets. With the addition of water jackets, the plastic target has a representative passive IR signature of an actual T-72 main battle tank. The results of the simulations presented here indicate that water jackets are a successful approach to improving the passive IR signature of the existing plastic ground target. Initial success in fabricating the water jackets and support structures indicate that the process is technically feasible. The simulation results verify the basic concept of using components with equivalent heat capacities for surrogate vehicle construction. With the addition of water jackets, the need for targets that can accurately simulate the visual and infrared signatures of threat systems has been met. The augmented plastic targets are fully capable of emulating infrared signatures of threat systems for use in destructive and non-destructive testing by the U.S. Army T&E community and there are vast opportunities for the usage of such targets.

REFERENCES

[1] Sanders, J.S., and Wible, W., "A Novel Method for Enhancing the Passive Infrared Signature of Target Surrogates", *Proceedings of the Sixth Annual Ground Target Modeling & Validation Conference*, Houghton, MI, August 1995.

Tracker Analysis and Ground Truth Tool Description
for the Proceedings of the 2002
Ground Target Modeling and Validation Conference

Carrie Kimbel, Jeff Todd, Mark A. Chambliss, Judson R. Griffin III,
Daniel Konkle, Paul D. Lavallee, Dr. Jay Lightfoot
Dynetics, Inc.
Huntsville, AL 35814 USA

ABSTRACT

The Tracker Analysis and Ground-Truth (TAG) tool was developed to streamline the process of defining a target region for each image in an image sequence in order to calculate signature metrics and perform tracker analysis. Previously, images within a sequence were analyzed individually by outlining the target in each image and calculating the relevant statistics. Likewise, tracker performance was determined by a man-in-the-loop viewing of tracker results overlaid on the image sequence. While effective, these methods did not allow the timely processing of long sequences and limited the calculation of signature metrics to only a few frames of a sequence. Using the TAG tool, a single rectangle is drawn around the target at the beginning of the sequence. The rectangle is shifted and resized using mouse motion and keystrokes as the sequence is played. The rectangle size and location is stored in a file which can be replayed and modified. This file can be used as input to SEMIRS (Software for Extracting Metrics from Infrared Sequences) for generation of signature metrics for each image and as stabilization input to ISAT (Infrared Seeker Algorithm Tool) which generates tracker output. The TAG tool has made it possible to generate, store, and utilize valuable ground truth information for entire sequences in a short period of time. This paper will focus on the utilization, functionality and application of the TAG tool.

INTRODUCTION

The process of analyzing infrared test imagery often requires ground-truth information, or the location and size of a target region. Ground-truth information is required to calculate signature metrics and to evaluate the performance of tracker algorithms. The target location and size define the target pixels that are to be used in metric calculations, and quantify tracker performance by determining the overlap between the target region and the track gate.

Prior to the creation of the TAG Tool, several commercial software packages were used in a piecewise fashion to generate the ground-truth data. First, image sequences were broken into individual image files in order to be imported in Scion Image (a.k.a. NIH Image). Each image was opened separately in Scion Image and an outline was drawn around the target. The pixels inside the outline were set to a value of two and the pixels outside the outline were set to zero. The modified image was then saved as the ground-truth file for that image. When complete, individual ground-truth files existed for every image in the sequence. It should be obvious that this method was very inefficient.

The TAG Tool was developed to streamline the process of generating ground-truth information by processing image sequences as a unit. To generate ground-truth information, a single rectangle is drawn around the target at the beginning of the sequence. Once outlined, ground-truth mode is invoked and the rectangle is resized and repositioned on every frame as the sequence is played. The rectangle can be resized using keystrokes and moved using mouse motion while in this mode. The location and size of the box are recorded for each frame and may be saved in a single text file. In this manner, ground-truth information for an entire sequence may be generated in a reasonable amount of time. For a sequence of about 3000 frames, it takes the typical user about 2 hours, while the manual method would take about 40 hours.

OPENING AN IMAGE SEQUENCE

In order to ground-truth an image sequence, the sequence must first be opened. The TAG Tool will open several formats: ADIF (Air Defense Image Format), NDIF (New Air Defense Image Format), bin

(Raytheon format), or raw. The raw format is especially useful because it is user-defined. A file may be opened using the 'Open ...' or 'Import ...' command under the 'File' menu as shown in ***Figure 1***, or using

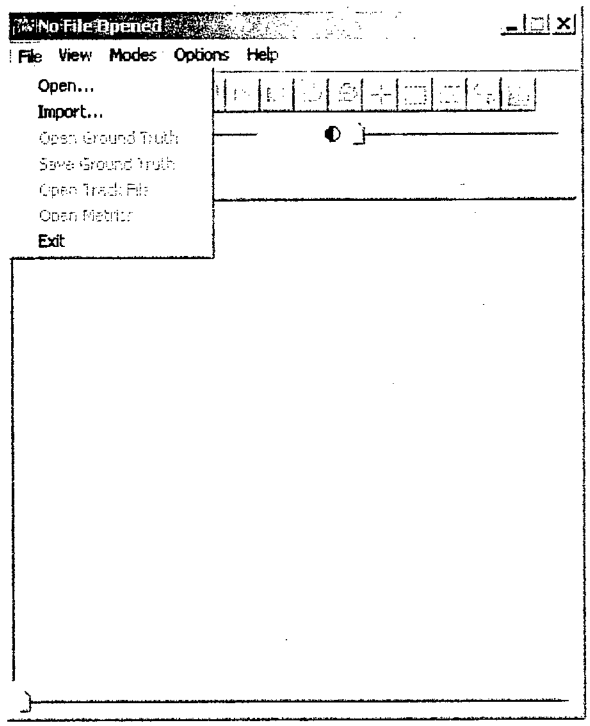


Figure 1. File Menu

toolbar button with the file folder icon. The 'Import File' command should only be used for raw image types. It will automatically bring up the raw image properties dialog box once a file is chosen. Files may also be opened by simply dragging and dropping them into the image display area. In this case, the file type will be determined using the file's extension. If the extension does not match a known format, it will be opened as a raw format. Once a file is opened, the name of the file will be displayed in the title bar.

PLAYBACK CONTROLS

The toolbar has seven VCR-like playback control buttons that control the playback of the image sequence. From left to right they are: play backward, rewind, step backward, step forward, go to frame, play forward, and stop. The 'go to frame' button will bring up a dialog asking for the destination frame number. If a specific frame number is not needed, but only a general location within the sequence, the slider bar underneath the image may be used to quickly advance to a general location within the sequence.

Keystrokes may also be used to play through the sequence. The stand-alone arrow keys are used for this and are applied as follows: up arrow = play forward, down arrow =stop, right arrow = step forward, left arrow = step backward. Play advance may be set to 1,2, 5, 10, or 20 frames per step under the 'Options' menu. This allows for a quicker viewing of the entire sequence.

The current frame number is displayed directly above the right hand side of the image. By default the frames are displayed counting down , starting with the total number of frames. This order may be reversed by using the 'Options' pulldown menu and selecting 'Count Up' under 'Image Numbers.' Two other pieces of information are also displayed directly above the image: the value in the center is the current zoom factor, and the values on the right hand side are the mouse location on the image and the corresponding pixel value.

The tool automatically scales the image for optimum viewing, but the user may adjust the brightness and contrast using the slider bars located directly under the toolbar. The brightness and contrast may be reset to their default value using the toolbar button with the brightness and contrast icons. The default brightness and contrast values for the initial frame of the sequence will be used for every image in the sequence unless the 'Automatic Gain Control' option is checked under the 'Options' menu. If this option is checked, the optimum brightness and contrast values will be recalculated for each frame that is played. Calculating these values for every frame will slow down the playback process, but depending on the speed of the computer, it may not be noticeable.

The image may be zoomed and re-centered using the toolbar buttons with the magnifying glasses and the crosshairs respectively. All of these buttons persist until they are pushed again or until another button is pushed. This persistence allows the user to zoom in or out or re-center as many times as needed without having to push the toolbar button each time.

GROUND-TRUTHING PROCESS

Ground-truthing may be done in forward or reverse mode, which means starting either at the beginning or at the end of the sequence respectively. The ground-truthing direction may be set under the 'Options' menu as shown in ***Figure 2***. Reverse mode is usually preferred since the target is easier to outline at the end of a sequence. The outline may then be adjusted as the sequence is played backward.

Once an image is opened, the target should be outlined. For reverse ground-truth mode, the target should be outlined on the last frame, or for forward ground-truth mode, the target should be outlined on the first frame. The outlining is done using the toolbar button that has an icon of a dotted outline of a rectangle. Once this button is clicked, the outline may be drawn around the target. Once drawn, the rectangle may be resized and moved using the mouse if adjustments are needed.

Since the target box will need to be adjusted between frames as the sequence is played, it may be necessary to slow down the playback speed. This can be done by adjusting the play timer, which is under the 'Options' menu, to the desired rate. The available playback rates are 10, 100, 500, 1000, and 2000 msec/frame. A good rule of thumb is to start out at the slowest rate and gradually increase the rate as needed.

Ground-truth mode must be selected for the target box to be recorded while the sequence is played. This mode may be activated by checking it under the 'Modes' menu as shown in ***Figure 3***. Once this mode is turned on, the target box will be recorded every time the frame is advanced in the direction in which the ground-truth direction is set. If the ground-truth direction is reverse, the target box will only be recorded when the sequence is played in reverse and when the ground-truth direction is forward, the target box will only be recorded when the sequence is played forward.

Keystrokes should be used to advance the frames while in ground-truth mode because mouse motion moves the target box while in this mode. The up arrow on the stand-alone arrow pad which causes the

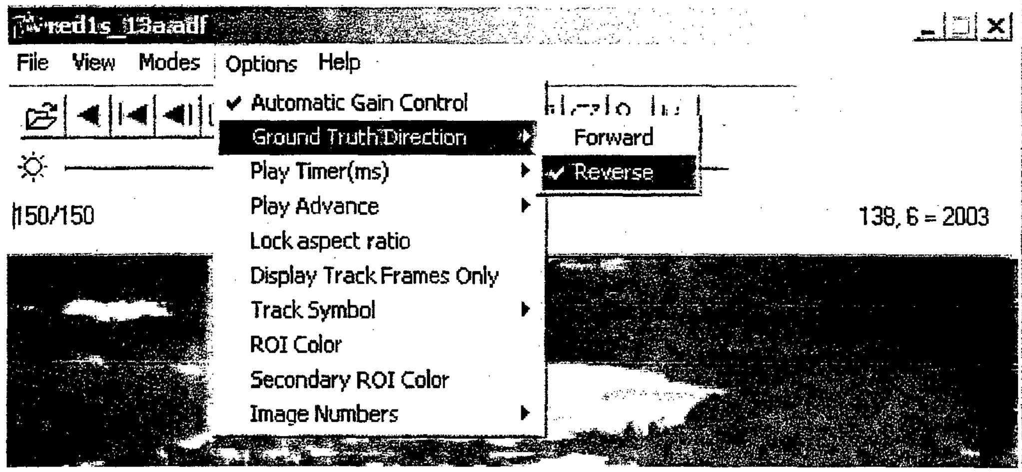


Figure 2. Options Menu

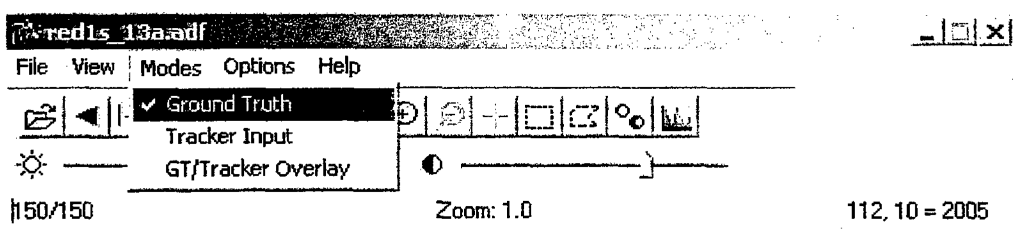


Figure 3. Modes Menu

sequence to play, will cause the sequence to play backward if reverse ground-truth mode is in effect. As the sequence plays and the target box is recorded, the box may be resized using the arrow keys on the number pad. The up and down arrows cause the height of the box to increase and decrease respectively, and the right and left arrows cause the width of the box to increase and decrease respectively. The aspect ratio of the original target box may be maintained automatically by checking the 'Lock aspect ratio' under the 'Options' menu. If this option is selected, the width will be automatically adjusted to maintain the original aspect ratio when the user adjusts the height and the height will be automatically adjusted when the user adjusts the width.

Once all the frames in the sequence have been ground-truthed, the target box information should be saved. This may be done by selecting the 'Save Ground Truth' command under the 'File' menu. Ground-truth information is usually saved with a 'gtr' file extension. The tool will choose a default filename that will be the image filename with a 'gtr' extension. This filename may be accepted, or the user may choose any filename allowable under Windows. For long sequences, it is recommended that information be saved periodically throughout the ground-truthing process.

REPLAYING A GROUND-TRUTH OR TRACKER FILE

Once a ground-truth file is created, the user may wish to view the target box information that has been saved. To do this, the file should be opened using the 'Open tracker file' under the 'File' menu. The default file extension for this operation is 'trk', but the file filter may be changed to see 'gtr' files. A tracker file may also be opened by dragging and dropping it into the image display area. However, it must have a 'trk' extension to be opened as a tracker file using this method. Next, the 'Tracker Input' mode should be checked under the 'Modes' menu. If this mode is selected and a tracker file has not been opened, the 'Open' dialog box will appear, and the user must select a file before the mode becomes active.

Once this mode is selected, a target box should appear over the image as long as the tracker or ground-truth file contained information for the current frame. As the sequence is played in this mode, the target box corresponding to each frame will be overlaid onto the image. This feature allows the user to check his or her work to make sure the target box was recorded as intended.

This mode is often used, and was designed, to view tracker simulation results. The tracker simulation output files must have a similar format to ground-truth files as described in the next section. Often tracker simulations do not process every frame, so when the file is played back, the target box appears to be blinking because it only appears on the processed frames. This effect can be bypassed by checking the 'Display Track Frames Only' item under the 'Options' menu. This option checks to make sure track information is available for a frame before playing it, and if not available, the frame is skipped.

There is also a GT (ground-truth)/Tracker Overlay mode, which may be used to compare two ground-truth and/or tracker files. The two files used in this mode are called the "tracker file" and the "GT overlay file". Both of these files may be opened using their respective commands under the 'File' menu. If either of these files is not opened before selecting the 'GT/Tracker Overlay Mode' command under the 'Modes' menu, 'Open' dialog boxes will appear to select them before this mode becomes active.

The symbol displayed for the track box may be changed to a crosshair instead of a box, or both box and crosshairs may be displayed simultaneously. This setting may be changed using the 'Track Symbol' item under the 'Options' menu.

GROUND-TRUTH AND TRACKER FILE FORMATS

Ground-truth files saved using the TAG tool use a simple tab delimited ASCII format. The first three lines are header lines containing the following information:

Line 1 – Corresponding image file, date and time ground truth file was saved
Line 2 – TAG version number
Line 3 – Column headings

All remaining lines are ground-truth information in tab delimited format with the following columns: frame number, top left row coordinate, top left column coordinate, height, width, slant range, row velocity, and column velocity. The slant range is extracted from the image sequence if it is available. The row and column velocities are generated based on shift in the ground truth box's center between the previous and current frame. The velocities are used as stabilization data in some tracker simulations.

Tracker files are expected to have a very similar format. They are also tab or space delimited ASCII files. Any line that does not begin with a number, or space and then a number, will be ignored. The first five columns of the remaining lines should have the same information as the ground-truth file. These columns are frame number, top left row coordinate, top left column coordinate, height and width. Any subsequent columns on a line will be ignored. It should be noted that the ground-truth file format conforms to the tracker file format. Therefore, when using tracker input mode, a ground-truth file may be used as instructed in the previous section.

MODIFYING A GROUND-TRUTH FILE

Sometimes it may be necessary to save a ground-truth file before the ground-truthing process is complete and finish it later, or to modify an existing ground-truth file to fix an error. First, the image numbers for which modification are required should be determined. The image file should be opened and advanced to one of the frames to be modified, or if finishing a ground-truth process, it should be advanced to the next frame for which ground-truth information is required. The ground-truth file to be modified should then be opened using the 'Open ground-truth' command under the 'File' menu. Ground-truth mode should then be selected, or made active.

Once ground-truth mode is active, any advancement of the sequence in the ground-truth direction will cause the ground-truth information for the previous frame to be updated (i.e. if the ground-truth direction is set to 'reverse,' advancement in the forward direction will not record the target box). Updating occurs

whether or not ground-truth information for that frame already existed. Therefore, care should be taken so that previous work is not overwritten. The target box should be positioned at the desired location for the frame to be modified and the sequence should be advanced one frame in the ground-truth direction, or put into play if all remaining frames are to ground-truthed.

IMAGE INFORMATION DIALOGS

The image information dialog was designed to view information about the current image without cluttering the main window of the tool. This dialog may be opened by selecting 'Image Information' under the 'View' menu, shown in ***Figure 4***. Initially, this dialog box will only show Irig time and slant range. More items may be viewed by selecting them under the 'View' menu of the image information dialog box as shown in ***Figure 5***.

The following metrics may be viewed if a metrics file is provided: GLCM (gray-level co-occurrence matrix) target, GLCM feature, GLCM total, DeltaT, DeltaTrss, and Scrrss. A metrics file usually has a 'txt' extension. It should be an ASCII text file in tab-delimited format containing: frame number, GLCM target, GLCM feature, GLCM total, DeltaT, DeltaTrss, and Scrrss. A column heading line or other header lines may be present in the file. Any line that does not start with a number will be ignored. A metrics file may be opened using the 'Open metrics' command under the 'File' menu in either the image information window or the main window, or it may be dragged and dropped into the display window if it has the 'txt'

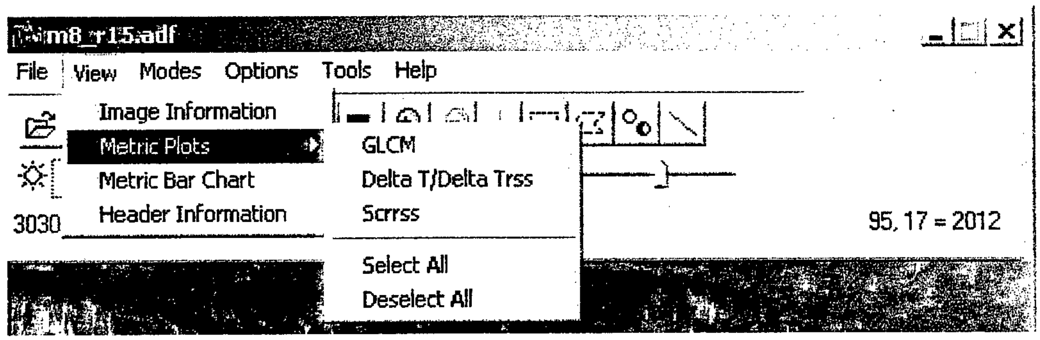


Figure 4. View Menu

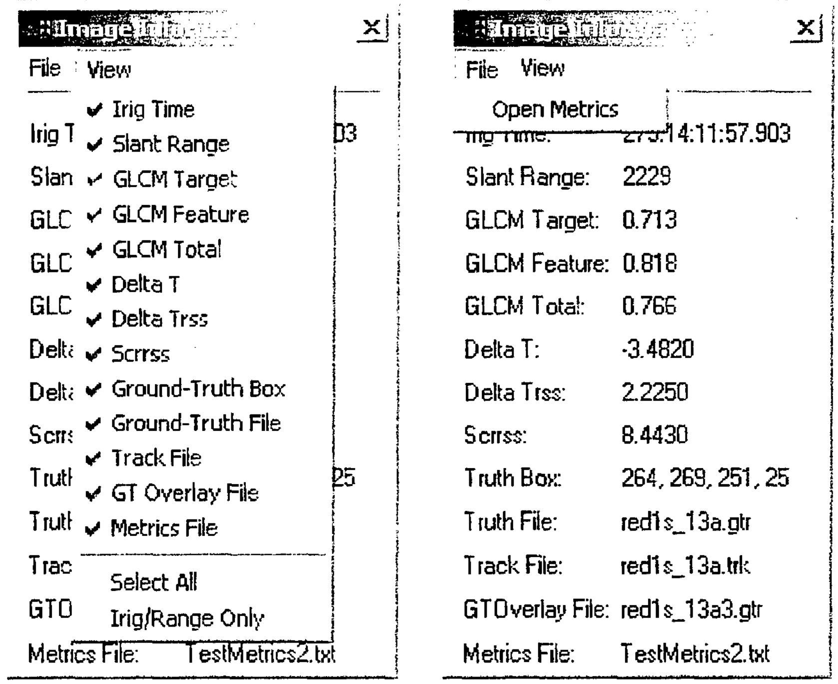


Figure 5. Image Information

extension. Other information items that may be viewed in this window are ground-truth box status, ground-truth file, tracker file, ground-truth overlay file, and metrics file. The ground-truth box status lists the coordinates of the ground-truth box for the current frame. If no ground-truth box exists for the current frame, then 'none' will be displayed. All available items may be displayed by choosing the 'Select All' command under the 'View' menu, as shown in ***Figure 5***, or the default of only Irig time and slant range may be displayed by selecting 'Irig/Range Only.'

A more basic form of image information is available under the 'View' menu by using the 'Header Information' command. This shows a simple list of header information items and the corresponding value for the item on that list which is selected. This information is only available for the ADIF and NDIF file formats. ***Figure 6*** shows the header information dialog box.

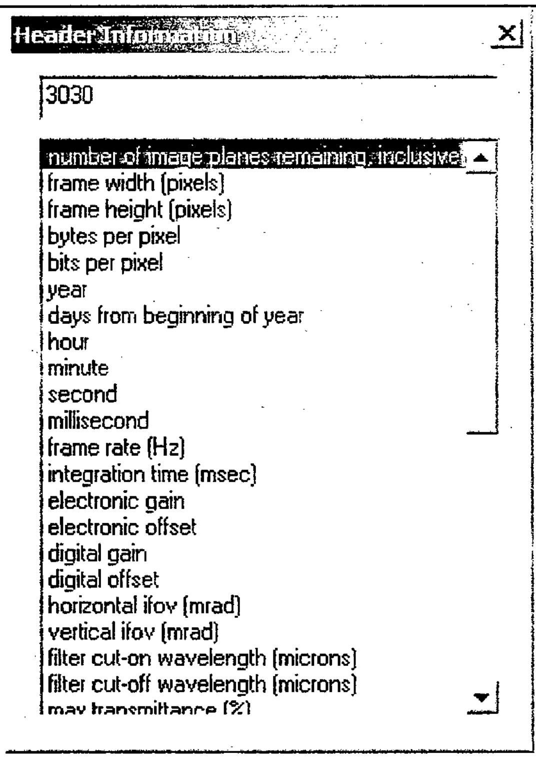


Figure 6. Header Information

METRIC PLOTS AND BAR CHART

In addition to the metric values displayed in the image information dialog box, metrics may also be displayed in plot and bar chart format. These charts may be accessed through the 'View' menu shown in ***Figure 4***. A metrics file must already be opened in order for these menu items to be enabled. Each plot may be selected individually through the 'Metric Plots' submenu, or they may all be selected or deselected as a unit using the two commands at the bottom of the submenu. ***Figure 7*** shows the bar chart and all three metric plots.

The bar chart is a good way to compare the different metrics for a single frame, while the metric plots are a good way to see the progression of a single metric throughout a sequence's history. The metric plots are like strip charts. They are generated as the sequence is played by adding a value to the right side of the plot each time a new frame is displayed and pushing values off the left side once the plot is filled. If the sequence is stopped, the plots are not cleared, so the values displayed on the plots do not necessarily represent sequential frames in the sequence. The plots are only cleared when they are closed, or when a new sequence is opened. The DeltaT/DeltaTrss plot shows two metric values since they are very closely related. Delta T is shown in red and DeltaTrss is shown in blue.

TOOLS MENU

There are four tools available under the 'Tools' menu, which is shown in ***Figure 8***: 'Histogram', 'Interpolate Slant Ranges', 'Interpolate Ground Truth', and 'Blackbody Calibration'. There is a fifth menu item, 'Record Avi', that has not yet been implemented.

A histogram of a rectangular area may be viewed by choosing the 'Histogram' tool. A rectangle must first be outlined for the histogram to be displayed. If a rectangle is not selected, an error message will appear. The limits of the histogram plot will be the minimum and maximum values of the rectangular area. More or less detail may be shown by changing the number of bins or the bin width. When one of these values is

changed, the other value is automatically calculated and changed accordingly. ***Figure 9*** shows an example histogram.

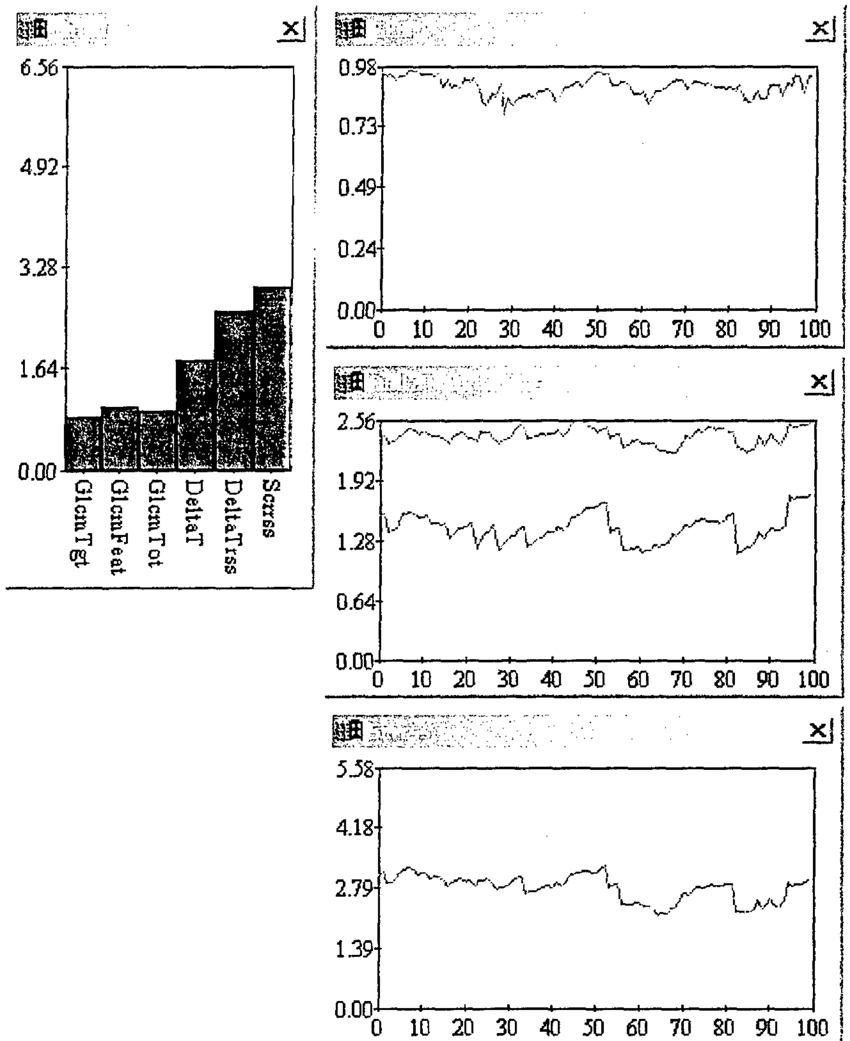


Figure 7. Metric Plots and Bar Chart

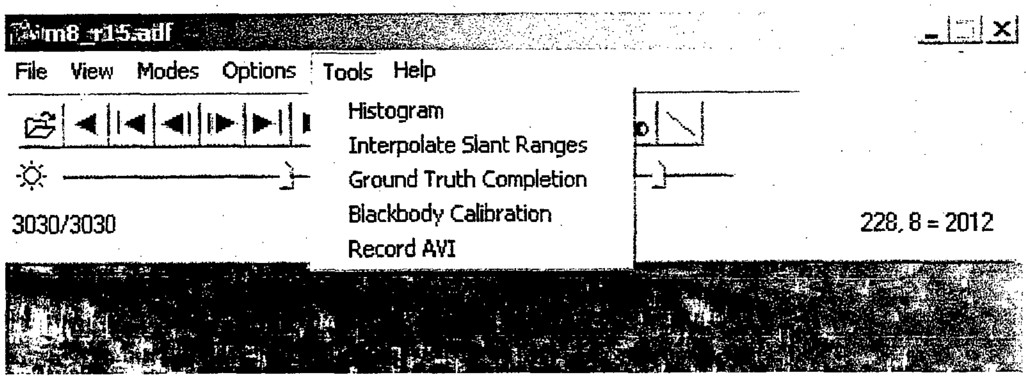


Figure 8. Tools Menu

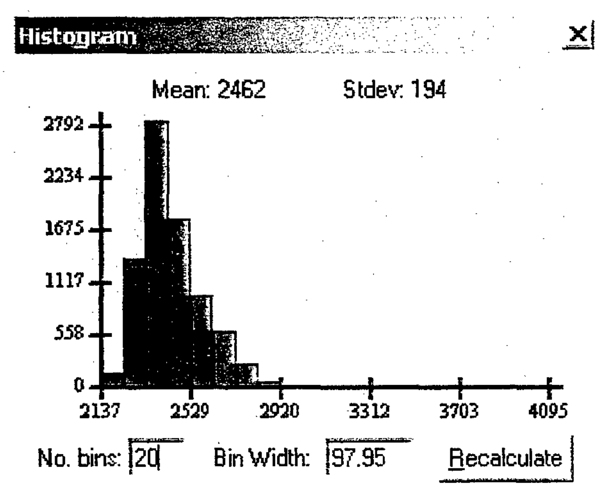


Figure 9. Histogram Tool

The 'Interpolate Slant Ranges' tool was added in conjunction with the 'Auto-size ROI' option. In order to facilitate accurate auto-resizing of the target box, the slant range information often needs to be smoothed. For example, if slant range information is updated once per second and the frame rate of the imagery is 30 Hz, then the slant range only changes for each 30 frames of imagery, giving a stair-step effect.. The tool simply performs a moving average over the existing slant range data and inserts the new slant range data into the sequence. The user must input the number of frames over which to perform each moving average calculation. For the example just given, the number to input should be 30, as shown in ***Figure 10***.

The next tool on the Tools menu is the 'Interpolate Ground-Truth' tool, shown in ***Figure 11***. This tool simply sets up a command line for a separate executable (GTR_Complete.exe). When the tool is run, a separate thread is created, so that the user may continue using the software while the ground-truth completion runs. Ground-truth completion is performed using a MADD correlation, using the previous and next ground-truth rectangles. The user must specify current and new filenames, which save the current and completed ground-truth information, and a desired search area (in pixels) for the MADD correlation.

In order to use the ground-truth completion tool, the currently loaded ground-truth must be evenly spaced. This ground-truth may be loaded using the 'Open ground-truth' command under the 'File' menu, or by generating the ground-truth and immediately using the completion tool. A study was done using both easy and hard to track targets, which showed that it is optimum to skip ten frames during the ground-truthing process if using the ground-truth completion tool. Skipping frames during the ground-truthing process is done by setting the 'Play Advance' under the 'Options' menu.

The last tool on the 'Tools' menu is the blackbody calibration tool, shown in ***Figure 12***. This tool will take up to ten calibration points, and will generate a 1st (linear) to a 6th order polynomial. Blackbody regions (ROI's) may be defined by outlining a region on the image with the rectangle tool and hitting the 'Get ROI' button, or may be manually entered by typing the top, bottom, left, and right coordinates of the region in a comma delimited format. Temperatures corresponding to each blackbody region should be entered in the second column. The number of blackbody points entered must exceed the order of fit desired for the resulting polynomial. Results may be saved once the calibration is performed. The output file will contain information about how the calibration coefficients were calculated, including the blackbody regions and temperatures, the image file used, and the frame used within the image sequence.

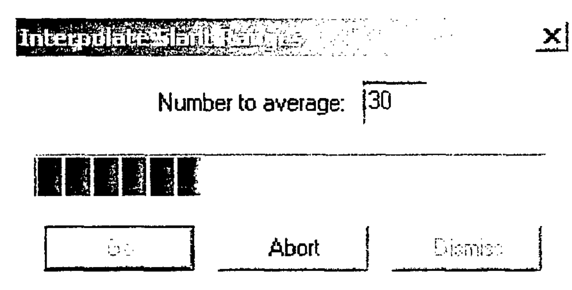


Figure 10. Interpolate Slant Ranges Tool

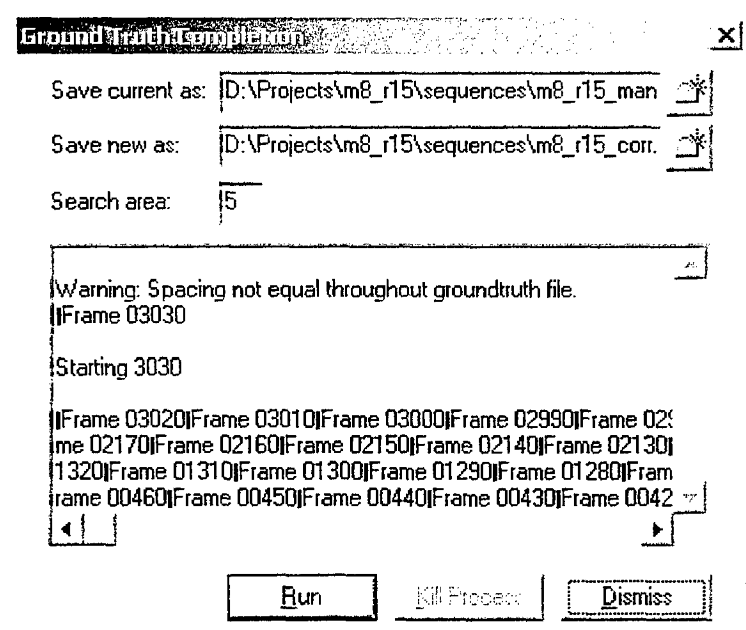


Figure 11. Ground-Truth Completion Tool

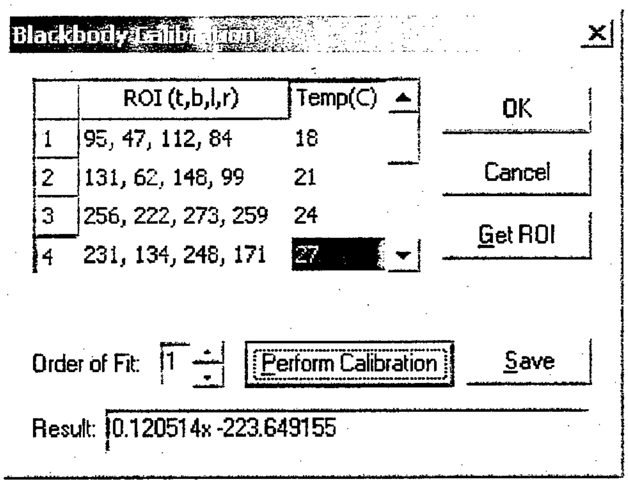


Figure 12. Blackbody Calibration Tool

ROI COLORS

The colors of the rectangular overlays, or ROIs (regions of interest) may be chosen by the user. This may be especially useful for the GT/Tracker Overlay mode to distinguish the track symbols. Users may choose the colors that they prefer at any time by selecting the 'ROI Color' or 'Secondary ROI Color' menu items under the 'Options' menu which was shown in ***Figure 2***. The colors will be stored in the Windows registry, so that they do not revert to the default every time the application is started.

HELP MENU

A few help items, as well as an 'About' box are available through the 'Help' menu shown in ***Figure 13***. The 'About' box, shown in ***Figure 14***, displays contact information for Dynetics, Inc, and gives acknowledgement to AMCOM (Aviation and Missile Command) for their sponsorship of the development of this application. The 'Help' dialog box, shown in ***Figure15***, currently contains four help items and more are planned as development of the software continues.

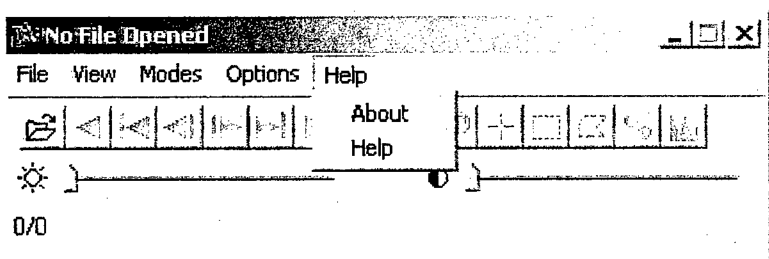


Figure 13. Help Menu

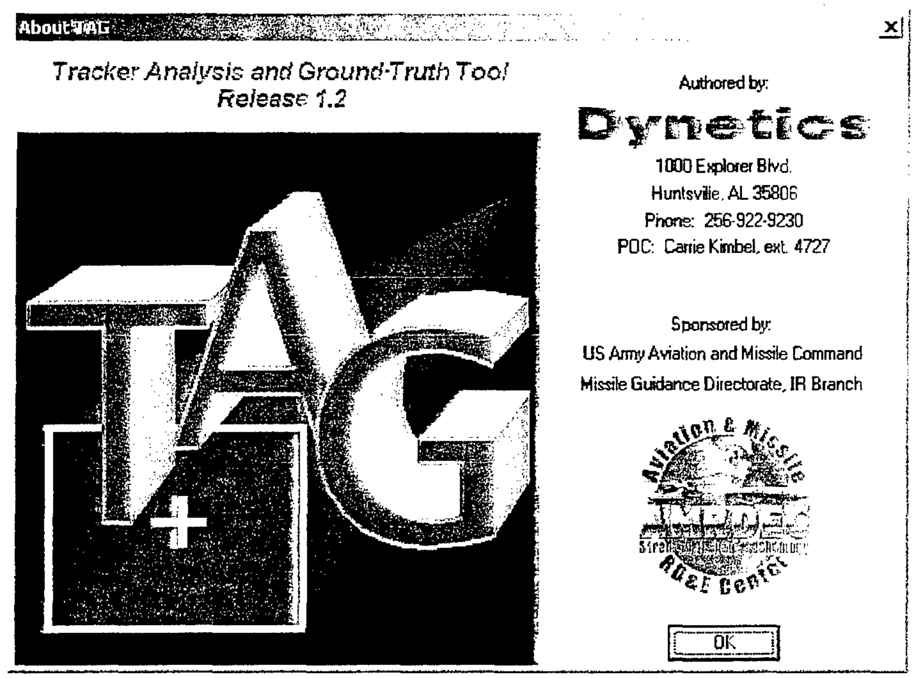


Figure 14. About Box

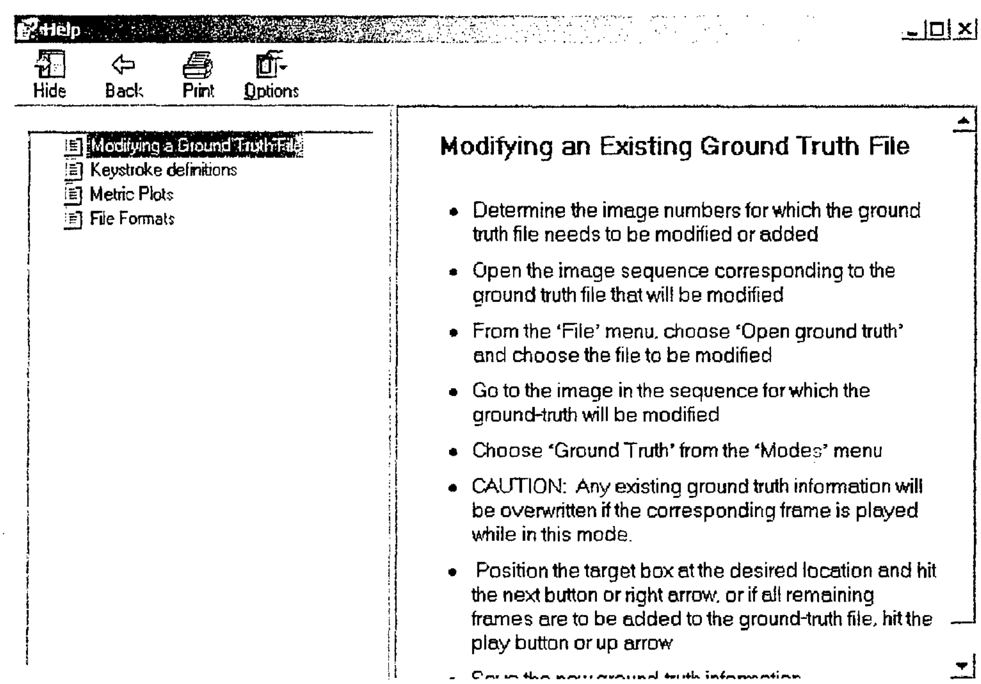


Figure 15. Help Box

CONCLUSION

The TAG tool has successfully streamlined the process of ground-truthing long image sequences for metric and tracker evaluation. During its development several features were created to make the process even easier, such as the ability to lock the aspect ratio of the ground-truth box, and the ability to ground-

truth in reverse mode. Features such as the tracker playback and ground-truth/tracker overlay make it a versatile tool that can be useful even for those not needing to generate ground-truth information. Future improvements to the tool are planned, including expanding the help information, adding a tool to create avi movies, and adding a tool to remove bad frames from a sequence.

ACKNOWLEDGEMENTS

This paper was based on work that Dynetics performed for the US AMRDEC, Missile Guidance Directorate, under the direction of Mr. Ron Passmore and Mrs. Lisa Cannon.

Benefits of Using the Photosimulation Laboratory Environment for Camouflage Assessment

Thomas J. Meitzler. David Bednarz, Euijung Sohn, Darryl Bryk, and Kimberly Lane
Survivability Technology Area
U.S. Army Tank-automotive and Armaments Command
Warren, MI, 48397-5000

ABSTRACT

A method is described for using the photosimulation laboratory environment to evaluate the effectiveness of camouflage for military vehicles. There are distinct advantages to acquiring images at the field site and then bringing them back for observer testing in a laboratory environment. Laboratory testing provides a repeatable, secure, and low-cost way to generate realistic performance data for vehicle evaluation for the purposes of signature testing, measurement of the effectiveness of camouflage relative to a baseline vehicle, and calibration and validation of target acquisition models. A test is described by the authors in which a baseline LAV is compared to a treated LAV in the TARDEC Visual Perception Laboratory using imagery collected from the field in the manner prescribed by an experimental design.

INTRODUCTION

High-resolution digital cameras presently available on the market are rapidly approaching the quality of film resolution. Five and six megabyte CCD imaging chips in combination with the ability to capture imagery in raw 24-bit format, combined with large memory storage devices enable high resolution imagery to be captured at field site locations. Using high-resolution graphics projectors, the imagery can be presented in the controlled environment of the lab in such a manner as to obtain observer data with confidence levels approaching 99%. It is the authors opinion that, the benefits of the high degree of confidence achieved using the repeatability offered by the lab environment, far outweighs any perceived advantage of having multiple teams of observers present at the field location to search for vehicles.

The Light Armored Vehicle (LAV) Family of Vehicles (FOV) was developed to provide the Marine Corps with enhanced mobile warfare capabilities. General Motors, Diesel Division in London, Ontario, Canada, began manufacturing the LAV FOV in 1982 and completed delivery to the USMC in April 1988. The LAV FOV includes several variants which utilize light armor protection from small arms, light machine gun fire, artillery projectile fragments, and mine fragments. Each variant was designed for a specific mission function and was mounted on a common chassis. The LAV FOV consists of direct fire variants armed with 25mm guns (LAV-25), with TOW II Missiles on LAV Anti-Tank (LAV-AT), an Air Defense variant (LAV-AD) capable of providing air defense coverage for the LAR unit, and support variants consisting of Command and Control (LAV-C2), Mortar (LAV-M), Logistics (LAV-L), and Recovery (LAV-R) vehicles.

The LAV is a highly mobile vehicle for conducting reconnaissance, counter reconnaissance, limited offensive and defensive operations and other missions. The current LAV is increasingly difficult to maintain at acceptable levels of readiness as it nears the end of its original projected service life. Budget constraints prevent the development and fielding of a new vehicle capable of performing its assigned missions. The main purpose of the Light Armored Vehicle Service Life Extension Program (LAV-SLEP) is to address the major subsystems that are currently degrading readiness or are driving higher than necessary O&S costs. The LAV retains the same requirements for performance capabilities and characteristics in the areas of mobility, lethality, survivability, and sustainability.

The test described in this report was undertaken as part of the LAV-SLEP. Specifically, the purpose of this field test was to determine the performance of a treatment known as ADCAM in reducing the probability of detection in the visual part of the electromagnetic spectrum at unity magnification (as seen by the naked eye) at various ranges, aspect angles and lighting conditions. Only the baseline results will be described in this report because comparisons of the treated versus untreated are classified.

METHOD

The test design implemented was an extension of visual detection requirements provided to our lab from PM LAV. Initially the PM requested only two range points at standard engagement ranges. We suggested having more ranges in between the critical range points to obtain a probability of detection versus range curve, as is more typical for these kinds of tests. A test matrix was developed in full-factorial form and 24-bit color imagery was collected using a Kodak 460 digital camera. The images were prepared for the photosimulation test and then presented to 30 subjects. The experimental factors and levels with their values are shown below in Table 1. The photosimulation test in the lab was arranged so that a pixel IFOV subtended by the display was less than one minute of arc and the displayed image represented a unity magnification or 1X representation to the subject. The first test was meant to emulate naked-eye vision. Prior to the actual test, the subjects were instructed on the purpose of the test as well as required to take a pre-test in which they could become familiar with the imagery and software. None of the pictures used in the pre-test were used in the actual test, however, the images were from the same set. The test procedure was to display an image with a time-out of 30 seconds. The imagery is cropped so that no scrolling is required. The target can appear within one of five possible regions. The soldier must use the mouse to "click-on" what her or she thinks is a target, based on the training.

Analysis of the first test showed most subjects obtained a score of only 20 % detection. This is not unreasonable given the difficulty of the imagery. The ranges are not unusual for such a test, however the high degree of clutter and in particular the height of the grass on the terrain makes it difficult for the unaided eye to detect common cure features of the vehicle. A second test was arranged at a power of 3X. The imagery from the field was of sufficient resolution so that there was no noticeable increase in pixelation of the imagery an in increase in magnification. The presentation in the lab was randomized, this is a very good reason to use the lab.

Region	
1	Top-Left
2	Top-Right
3	Lower-Left
4	Lower-Right
5	Center

Vehicle Type	
1	Baseline (old LAV)
2	SLEP + ADCAM
3	SLEP + ADCAM - ADCAM bowplane

Aspect angle	
1	Front
2	30 degree
3	Side

Lighting	
1	Front Lit
2	Back Lit

Weather condition	
1	Clear
2	Overcast

Range (km)	
1	1
2	1.5
3	2
4	2.5
5	3

Table 1: Factor matrices for the visual detection test

At this point, it's instructive to digress for a moment to emphasize the rationale for using the type of design methodology described by the authors in this paper. Statistically based experimental design is a strategy of designing experiments in such a manner as to develop a robust test plan. In other words, a test plan that is minimally affected by external sources of variability. What makes vehicle field test challenging is that there are many variables that are present and must be accounted for. In addition, the variables interact with each other. The correct approach to working with several factors is to conduct a factorial experiment. A factorial experiment is an experimental strategy in which the factors are varied together, rather than one at a time. The factorial experimental design concept is extremely important and powerful when used correctly. Entire books are written on how to choose an experimental design based on considerations such as time, money and availability of materials. Montgomery wrote a book that is a standard in the field of experimental design.[1] For the type of tests that the authors are involved with, it is dangerous to leave out data points in the hope of making the test plan more expedient. Attempts to do just that in the past have resulted in things such as the 'heartbeat effect' in data. The range of variability in the factors is so great and interdependent, that experimental designs such as fractional factorial designs are to be avoided.

When making inferences about differences in a factor in a perception experiment in the laboratory we want to make the experimental error as small possible. This requires that we remove the variability between subjects from the experimental error. The design we use to accomplish this is a factorial experiment run in a randomized complete block. By using this design with the subjects as blocks we form a more homogeneous experimental unit on which to compare different factors. This experimental design improves the accuracy of the comparisons among the different factors by eliminating the variability among the subjects. Within a block, the order in which the treatment combinations are run is randomly determined. It is usually not possible to implement this experimental design in the context of a traditional field test.

The pictures below in Fig. 1 through Fig. 6 were used for training observers as to what kind of vehicles they would be looking for.

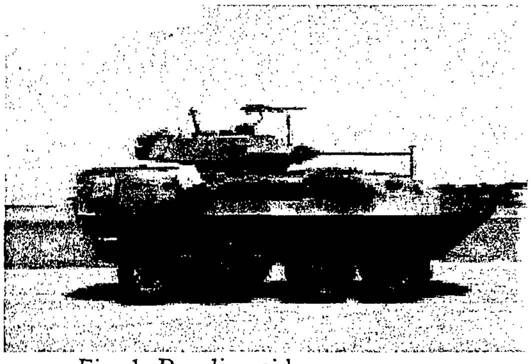


Fig. 1: Baseline side

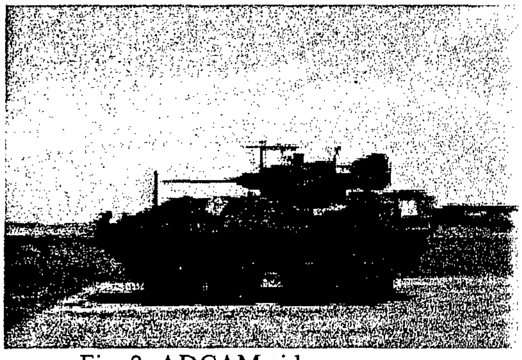


Fig. 2: ADCAM side

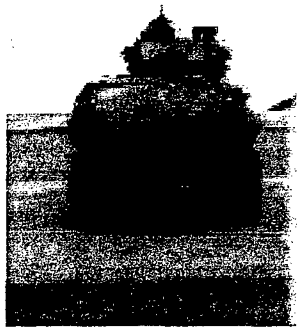


Fig. 3: Baseline front

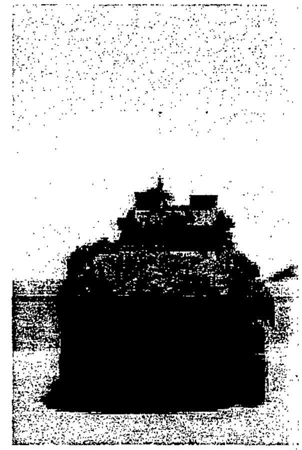


Fig. 4: ADCAM front

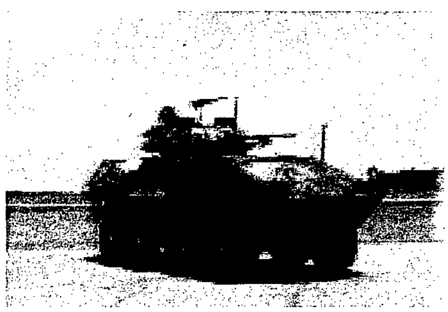


Fig. 5: Baseline at 30 degrees

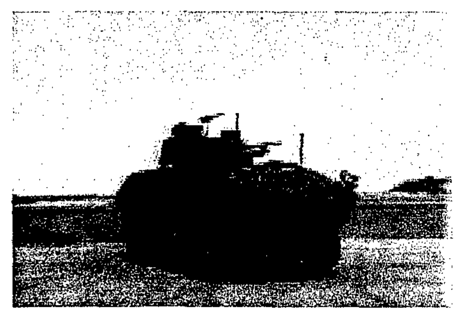


Fig. 6: ADCAM at 30 degrees

The figure below is of the background at the field site and does not have a vehicle in it. The grass height was high at the test site.

Fig. 7: View of the test field

The charts in Fig. 8 below show the results of measuring the X and Y chromaticity values of the monitors that were used in the test. The values measured were compared to standard values and found to be virtually identical.

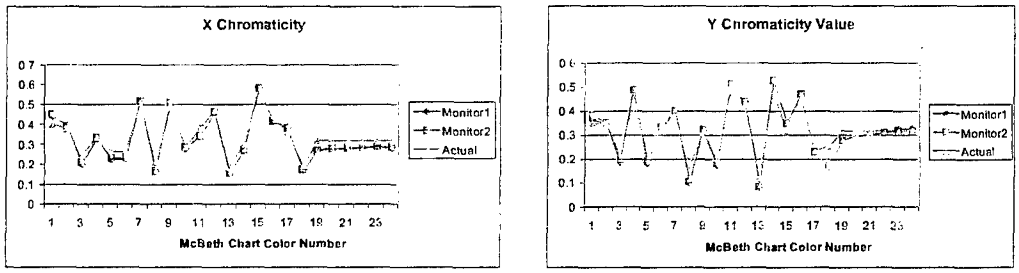


Figure 8: X and Y monitor chromaticity calibration charts

ANALYSIS

Below in Table 2 is the ANOVA table for the baseline vehicle and the other experimental factors. The treated vehicle has been excluded because of security classification. The power of the experimental design methodology is shown here in that the significance of individual factors and of their interactions are available. Using this kind of a test, one can obtain not only a model curve of the detection probability versus any factor in the test, but, one can also obtain the relative importance of the individual factors.

Tests of Between-Subjects Effects

Dependent Variable: RANK of RESPONSE

Source	Type IV Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	185215009[b]	89	2081067.514	26.782	.000	2383.607	1.000
Intercept	1159759015	1	1159759015	14925.408	.000	14925.408	1.000
SKY_COND	2918024.068	2	1459012.034	18.777	.000	37.553	1.000
RANGE	161301308	9	17922367.58	230.650	.000	2075.852	1.000
ASPECT	944347.896	2	472173.948	6.077	.002	12.153	.887
SKY_COND * RANGE	5751990.053	18	319555.003	4.112	.000	74.025	1.000
SKY_COND * ASPECT	2459473.720	4	614868.430	7.913	.000	31.652	.998
RANGE * ASPECT	6204854.010	18	344714.112	4.436	.000	79.853	1.000
SKY_COND * RANGE * ASPECT	5397720.861	36	149936.691	1.930	.001	69.465	1.000
Error	128521871	1654	77703.671				
Total	1641367780	1744					
Corrected Total	313736880	1743					

a. Computed using alpha = .05

b. R Squared = .590 (Adjusted R Squared = .568)

Table 2: ANOVA of test factors

Figure 9 shows the model generated logistic curve of the probability of detection of the baseline LAV. This curve has the effects of all the various factors 'rolled-up' into it.

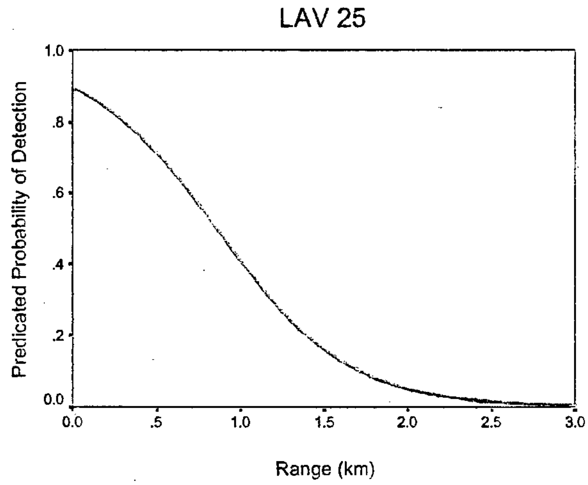


Fig. 9 : Logistic curve fit to the model from the subject responses

Fuzzy Logic Modeling of the DATA

The Fuzzy Logic Approach (FLA) was also used to model the experimental observer response detection data. The FLA and it's application to modeling the probability of detection is described in other papers by the authors. [2,3] The main elements of the model as applied to this test are shown below. The correlation obtained in this test was 0.9 between the experimental result and the FLA model predicted value. The 0.9 correlation is between the model built from half the data set and half used as testing. Figures 10 to 13 show the several interfaces that are part of the FLA model and the resulting output surface.

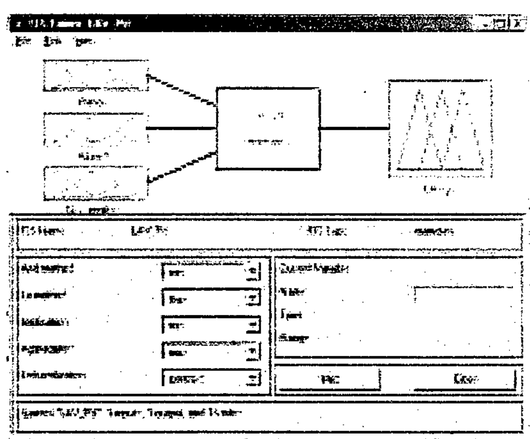


Fig. 10. FLA Fuzzy Inference Main Module

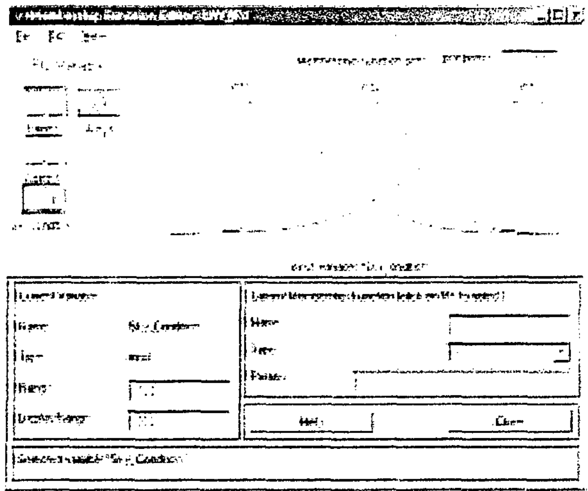


Fig. 11: FLA membership functions

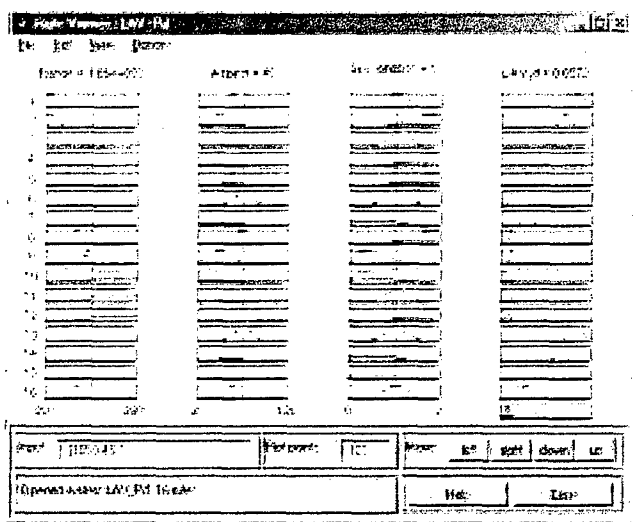


Fig. 12: FLA firing diagram

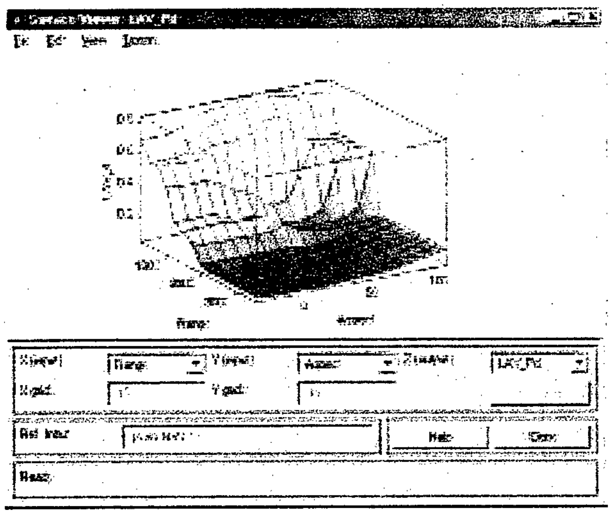


Fig. 13: Resulting model surface of Probability of detection versus range and aspect angle

Conclusions

An experimental design method was used to design an imagery collection test plan and a laboratory testing procedure. The probability of detection was determined for the baseline and treated vehicles. A statistical model was made of the laboratory results that gave probability of detection versus range. A fuzzy logic model was also made from the data that had 0.9 correlation to data not used in the training set.

An advantage of using the photosimulation lab environment is that we are able to archive scenes used in the simulation. Therefore, at a later time we are able to rerun the same test on a different subject pool. The new subjects may have a different training and the images may also be modified by either magnification or adding atmospheric conditions. This provides tremendous cost savings since we do not have to pay for another field test.

References

[1] Montgomery, C. Douglas, **Design and Analysis of Experiments**, Fourth Edition, John Wiley & Sons, New York, NY, 1997.

[2] Meitzler, T., Sohn, E., Singh, H., Elgarhi, A., and Nam, D., Predicting search time in visually cluttered scenes using the fuzzy logic approach," Opt. Eng., 40(9), Sept. 2001, pp. 1844-1851.

[3] Meitzler, T, Singh, H., Arefeh, L., Sohn, E., and Gerhart, G., "Predicting the Probability of target detection in static infrared and visual scenes using the fuzzy logic approach," Opt. Eng., Vol. 37 (1), Jan. 1998.

Improvements to the Gray-Level Co-occurrence Matrix (GLCM) Based Trackability Metric

Ricky K. Hammon
Missile Guidance Directorate
Missile Research, Development. and Engineering Center
US Army Aviation and Missile Command. Redstone Arsenal, Alabama

Mark A. Strauser
Computer Sciences Corporation,
Huntsville, Alabama

ABSTRACT

The Gray-Level Co-occurrence Matrix (GLCM) based Trackability Metric has bee used at AMCOM for many years for autotracker performance evaluation. The origins of the metric stem from the frustrating experience of trying to use such measures as delta T and signal to noise ratios to specify autotracker performance requirements. This paper presents the most recent developments and improvements to the GLCM based Trackability Metric (TM) and the resulting performance enhancements. One of the new developments is the inclusion of a hot spot or most predominant feature metric. Additionally, the results of a sensitivity study for the accuracy of target segmentation verses the GLCM TM performance is presented. These new developments for the Trackability Metric should provide better state-of-the-art performance prediction and more accurate performance modeling for specific imaging autotracker designs. The results of these studies and the final implementation of the Trackability Metric are presented.

Keywords: imaging, feature, resolved, target, signature, tracking, clutter, metric, autotracker

1. INTRODUCTION

The GLCM Trackability Metric (TM) was originally developed by the U.S. Army Aviation and Missile Research, Development, and Engineering Center (AMRDEC) under the Advanced Imaging Autotrackers Technology Base Program to predict the performance of autotrackers for imaging infrared (IR) missile seekers. The original application was to predict how well an autotracker could successfully segment and track a target in its entirety from the background. For some applications full segmentation is important, for other applications it is not. We found that when the TM, as originally designed, was used for applications without a full segmentation requirement, the TM often did not correctly predict the resulting autotracker performance. Some targets with very low full segmentation TM values were easily tracked. This was sometimes caused by a single small predominant feature on the target, which allowed the autotracker to track the single feature itself against the combination of the remainder of the target and the background.

Several variations of the TM were evaluated to include the effects of a predominant target feature on autotracker performance. Several feature segmentation and predominance measurement algorithms, and methods of including this new measurement in the overall TM representation, were studied. Only the resulting design is presented here in this paper.

User feedback indicates that the time intensive ground-truthing process required for the application of the TM has lead to short cuts in the target segmentation process. The obvious question then becomes how accurate does the target segmentation have to be and what is the impact on the accuracy of the TM? The results of a study addressing this issue are addressed in section 6.

2. GLCM Trackability Metric Revisited

The GLCM is a square matrix whose elements are an estimate of the joint probability that a pair of pixels will have two specific grayscale values. The formal definition of the GLCM, found over a region of interest (ROI) within an image and for a given orientation, is given by Equation 2-1.

$$G_{a,b}(r_\theta,\theta) = \frac{N_{a,b}}{N} \approx \Pr\{I(m,n)=a, I(m+r_\theta\cos\theta, n+r_\theta\sin\theta)=b\}, \tag{2-1}$$

where

$I(m,n)$ = the image intensity or grayscale value at the m'th row and n'th column,

$N_{a,b}$ = the number paired pixels whose grayscale values are a and b respectively,

N = the total number of pairs in the ROI,

r_θ = the quantized radial displacement between the pixels, and

θ = the polar angle or direction between the pixels.

The target and background GLCMs are found using appropriate windows as the reference Region of Interests (ROI)s for evaluation purposes. Pixel membership is determined using segmented-target ground-truth data for the image. The structure length of the target for a specified direction is given by locating the first peak in the minimum absolute difference (MAD) auto-correlation surface of the target along the direction of displacement. To account for noise and measurement uncertainty, the GLCMs are convolved with a kernel formed from the truncated noise distribution appropriate for the target area of interest. The same convolution kernel is used for both the target and background GLCMs and is used for four directional evaluations.

The measured TM value for the target area of interest is found by comparing the GLCM of the target with that of the local background using Equation 2-2. Equation 2-2 yields a directional TM value along a specified orientation. The final TM value is the average of four directional TM calculations evaluated at 0, 90, and ±45 degrees.

$$TM(\ell_\theta^{tgt},\theta) = 1 - \frac{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{tgt}(\ell_\theta^{tgt},\theta)\cdot\hat{G}_{a,b}^{bkg}(\ell_\theta^{tgt},\theta)\right]}{\sqrt{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{tgt}(\ell_\theta^{tgt},\theta)\right]^2}\cdot\sqrt{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{bkg}(\ell_\theta^{tgt},\theta)\right]^2}} \qquad (0 \le TM \le 1), \tag{2-2}$$

where

$\hat{G}_{a,b}^{tgt}$ = $G_{a,b}^{tgt}(\ell_\theta^{tgt},\theta) * K_{\Delta_a,\Delta_b}^{unc}(t)$, the convolved target GLCM,

$\hat{G}_{a,b}^{bkg}$ = $G_{a,b}^{bkg}(\ell_\theta^{tgt},\theta) * K_{\Delta_a,\Delta_b}^{unc}(t)$, the convolved background GLCM,

$G_{a,b}^{tgt}$ = the GLCM of the target area in the current frame,

$G_{a,b}^{bkg}$ = the GLCM of the local background in the current frame,

$K_{\Delta_a,\Delta_b}^{unc}$ = the truncated uncertainty kernel for the current frame,

ℓ_θ^{tgt} = the target's structure length along the current direction,

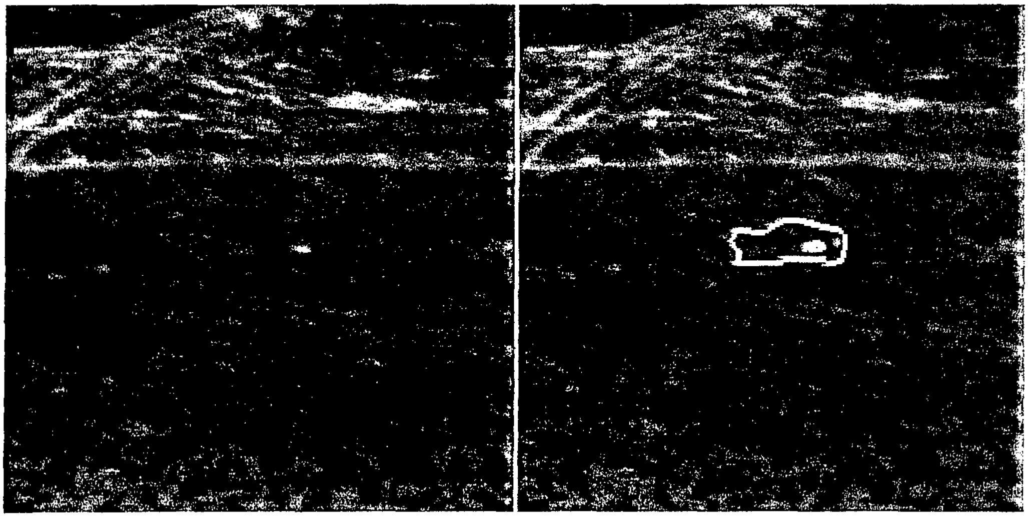
θ = the current direction of interest (one of the four possible values), and

L = the number of available grayscales in the image.

Notice that both the target and background GLCM's are found using the target's structure length. This emphatically couples the background clutter characterization with the target's structure. Note also that Equation 2-2 is still a directional quantity. Depending on the intended application, this can be exploited for comparing the target signature along different orientations or the orientation dependence can be integrated out by averaging the results over many directions. Further detail on the GLCM Trackability Metric fundamentals can be found in References 1, 4, and 5.

3. PROBLEM SCENARIO

The first change made to the TM was the introduction of a predominant feature measurement. This requirement has been added due to the application of the TM to new programs. One of the original requirements was to measure the likelihood of correctly segmenting the entire target from the background. Obviously this is not a requirement for a successful track for many applications. The particular case that drove the Predominant Feature Measurement is presented here. Figures 3-1 demonstrates typical IR imagery for which the GLCM TM measurement for the whole target is low but a relatively small but bright hot spot is present at the target's engine port. In this case, the number of pixels comprising the hot spot is small compared to that of the whole target; consequently, the hot spot makes a rather small contribution to GLCM TM of the whole target.



Original Image

Same Image with Target Mask and Feature Outlined

Figure 3-1. Low GLCM TM with Predominant Target Feature, Actual Data

4. DETERMINATION OF FEATURE PREDOMINANCE

A number of methods were explored to measure the level of target feature predominance that is present within an IR image. This measure would be combined with the classical whole-target GLCM to form a composite measurement. Included in these were the features relative SNR, the area ratio of the feature with the target, the energy ratio of the feature with the target, and many more elaborate formulations.

The selected implementation was to use the GLCM TM itself as a feature predominance measurement. This new implementation of the GLCM TM is the same as the classical form but with the feature area mask used instead of the whole-target, with the local background area is defined as anything but the feature. The result is a normalized form that

measures the textural content and the intensity of the feature relative to that of its background. In such cases as shown by Figure 4-1 and Equation 4-1 where a small but bright hot spot at an engine port is present, the extracted feature itself has a very different textural composition than that its background. This provides an excellent measure of an internal target component that can be used for tracking and, by the nature of Equation 4-1, is normalized.

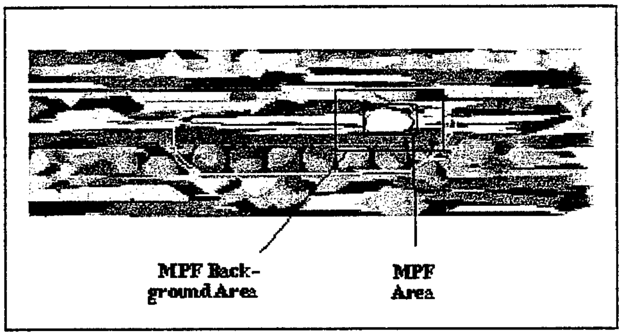


Figure 4-1. Feature GLCM Trackability Metric, Measure of Feature Predominance

$$TM_{MPF}\left(\ell_\theta^{MPF},\theta\right) = 1-\frac{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{MPF}\left(\ell_\theta^{MPF},\theta\right)\cdot\hat{G}_{a,b}^{bkg}\left(\ell_\theta^{MPF},\theta\right)\right]}{\sqrt{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{MPF}\left(\ell_\theta^{MPF},\theta\right)\right]^2}\cdot\sqrt{\sum_{a,b=0}^{L-1}\left[\hat{G}_{a,b}^{bkg}\left(\ell_\theta^{MPF},\theta\right)\right]^2}} \qquad \left(0\le TM_{MPF}\le 1\right), \qquad (4\text{-}1)$$

where

MPF	=	the Most Predominant Feature of the target,
$\hat{G}_{a,b}^{MPF}$	=	$G_{a,b}^{MPF}\left(\ell_\theta^{MPF},\theta\right)*K_{\Delta_a,\Delta_b}^{unc}(t)$, the convolved MPF GLCM,
$\hat{G}_{a,b}^{bkg}$	=	$G_{a,b}^{bkg}\left(\ell_\theta^{MPF},\theta\right)*K_{\Delta_a,\Delta_b}^{unc}(t)$, the convolved background GLCM,
$G_{a,b}^{MPF}$	=	the GLCM of the MPF area in the current frame,
$G_{a,b}^{bkg}$	=	the GLCM of the local background in the current frame,
$K_{\Delta_a,\Delta_b}^{unc}$	=	the truncated uncertainty kernel for the current frame,
ℓ_θ^{MPF}	=	the MPF's structure length along the current direction
Background	=	Every thing in the local area except the MPF.

5. COMPOSITE TRACKABILITY METRIC

A rational solution for a composite track metric (TM_{COMP}) is the normalized Root Sum Square (RSS) value of the whole-target track metric (TM_{tgt}), which is the classical form shown by Equation 2-2, and the feature-only track metric (TM_{MPF}) as is shown by Equation 4-5. This formulation takes into account the special problem scenario exhibited in section

3. When the GLCM-based TM algorithm is invoked using only the MPF as the target mask in such cases, the feature area is so much *unlike* the surrounding background that a relatively high value of TM_{MPF} is measured. With this type of case, TM_{COMP} would take into account the fact that the whole target would be difficult to track utilizing a multi-feature based tracker. It would also consider the fact that a portion of the target *could* be tracked but by ignoring the rest of the target. The normalized RSS combines these two measurements with the proper weighting, never allowing TM_{COMP} to be less than TM_{tgt}. In situations for which the whole-target structure very closely resembles the background and no predominant feature exists, TM_{COMP} would be measured as a relatively low value since TM_{MPF} would essentially be zero. The TM_{comp} formulation shown by Equations 5-1 and 5-2, and the properties of which shown by Figure 5-1 exhibit the nature of the composite metric. The *normalized* RSS seems to be the appropriate formulation for the low TM_{tgt} /hot spot condition as compared to a non-normalized RSS formulation, which was tested during algorithm development. The non-normalized RSS, limited to 1 does not achieve the objective since a very predominant hot spot and a low TM_{tgt} results in a saturated value with too much weighting on the feature component. More detail on the implementation and test results can be found in reference 6.

$$\text{For } TM_{MPF} > TM_{tgt}, \quad TM_{COMP} = \sqrt{1/2\left(TM_{tgt}^{\ 2} + TM_{MPF}^{\ 2}\right)} \quad (0 \le TM_{COMP} \le 1), \quad (5-1)$$

$$\text{For } TM_{MPF} <= TM_{tgt}, \quad TM_{COMP} = TM_{tgt} \quad (0 \le TM_{COMP} \le 1), \quad (5-2)$$

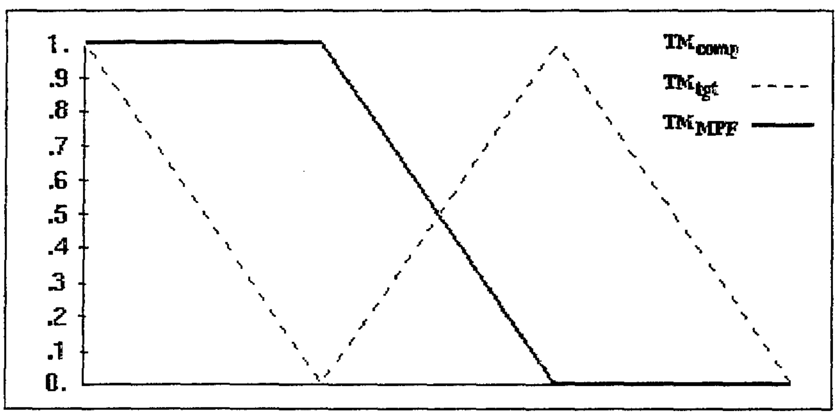


FIGURE 5-1. Composite GLCM Trackability Metric Properties

6. GLCM TM SENSITIVITY TO GROUNDTRUTH

Exact ground-truth is available for synthetic image sequences, but is not available for captive flight data. In order to analyze tracker performance against captive flight data, the ground-truth data must be generated by hand. This manually generated ground-truth data is obviously not as accurate as it would be if it had come from a simulation. The first questions become, how does human error in the ground-truth data affect the TM and can one ground-truth an image sequence well enough to get reasonable TM values? Once we have these answers, we must develop and provided a tool and technique for ground-truthing that reduces the workload enough to make the TM a usable product.

The ground-truth can have many sources of inaccuracy. Some of the sources noted in the past have been the inclusion of background pixels in the target definition, the omission of some of the target pixels from the target definition, and target definitions consisting of a rectangular box approximation of the target instead of an accurate segmentation. These inaccuracies are always a direct result of the labor and time intensive Groundtruthing process. Shortcuts are required to allow the TM to become a cost effective tool.

7.0 TRACKLAB GROUND-TRUTHING TOOL

The ground-truthing capabilities of TrackLab were enhanced with the inclusion of three-dimensional models of a T-72 main battle tank and a BMP2 armored personnel carrier. This greatly reduces the amount of time required to ground-truth captive flight test (CFT) data. Keyboard commands where added to TrackLab to control the position, scaling, and rotation of the models. Figure 7-1 shows a MICOM IR Seeker Analysis Tool (MIRSAT) seeker image with an overlay of the 3-D target model.

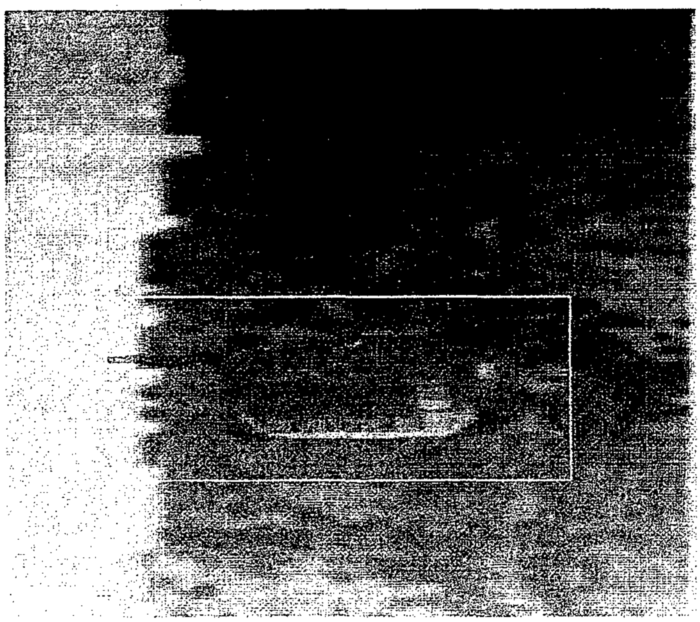


Figure 7-1. 3-D Target Model Overlay.

The position, scaling, and rotation values for the model are stored as a keyframe and are interpolated in between keyframes. The user then has the option of initiating a Minimum Absolute Difference (MAD) correlation to accommodate minor jumps in the imagery. A new reference image is captured every five frames, or whenever a keyframe is reached.

With the assistance of the TrackLab tool, an experienced user can ground-truth an image sequence in a respectable amount of time. The image sequence represented by the image in Figure 7-1 contained 1150 images and was ground-truthed in approximately one hour. There were 81 keyframes defined during the process. The operator identified the key frames when there was visible error in the target model placement and the target in the image. The operator, while trying to do a good job, was instructed to "hurry" in order to reduce the time required for Groundtruthing as much as possible. The tradeoff for accurate TM measurements is time and money for the ground-truthing process.

TM Groundtruthing Trade Study

This study was undertaken to determine how sensitive the GLCM TM is to the accuracy of the ground-truth data. The first step in this study was to use a synthetic sequence, which included "perfect" ground-truth data. The rows of ground-truth data were then shifted up by one, two, and three rows. The GLCM TM was then run using the original ground-truth as well as the shifted versions. Figure 7-2 shows the results.

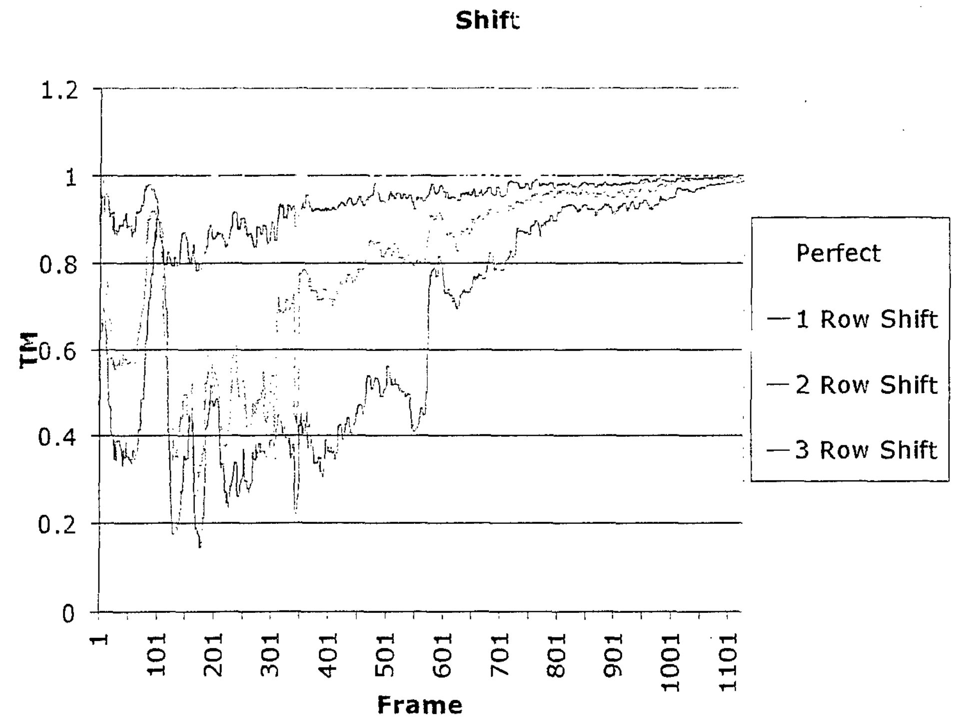


Figure 7-2.

The output displayed has been filtered to make the chart more readable. Figure 7-3 shows the percent errors of the shifted curves with respect to the perfect one. The curves are just what would be expected. At the start of the sequence, when the target is small, minor errors have major effects. And near the end of the sequence, the target is much larger taking up several rows of the image. Therefore, minor errors in the ground-truth placement have little effect.

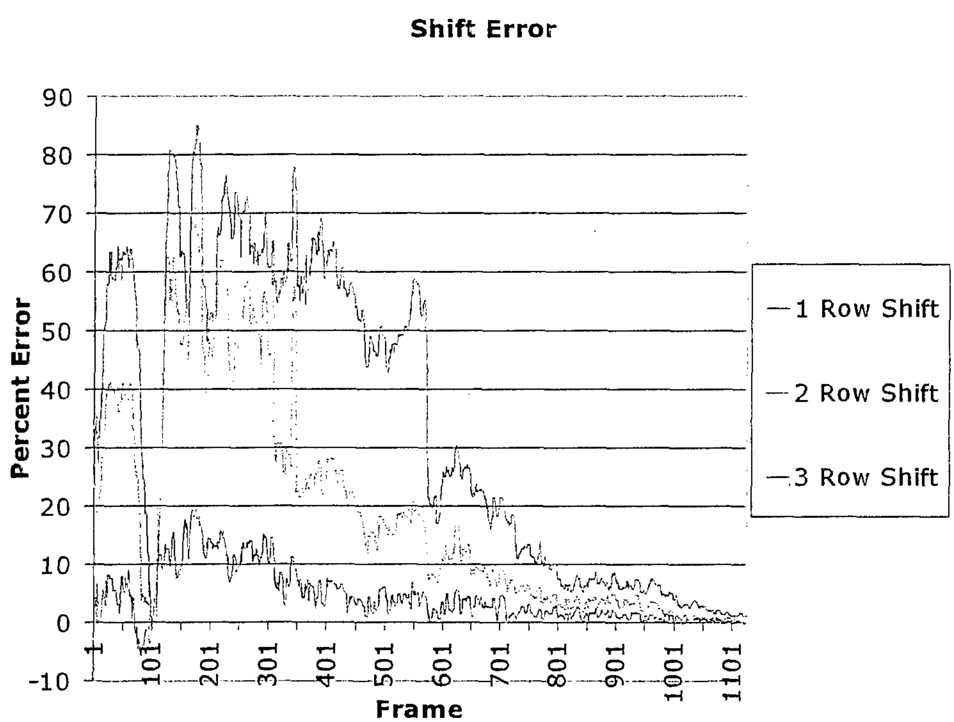


Figure 7-3.

The next step of the study was to use TrackLab to manually ground-truth the synthetic imagery. This was done in a "quick and dirty" manner to give worst-case results. Figure 7-4 shows the TM values calculated using the perfect and manual ground-truth data and Figure 7-5 shows the percent error.

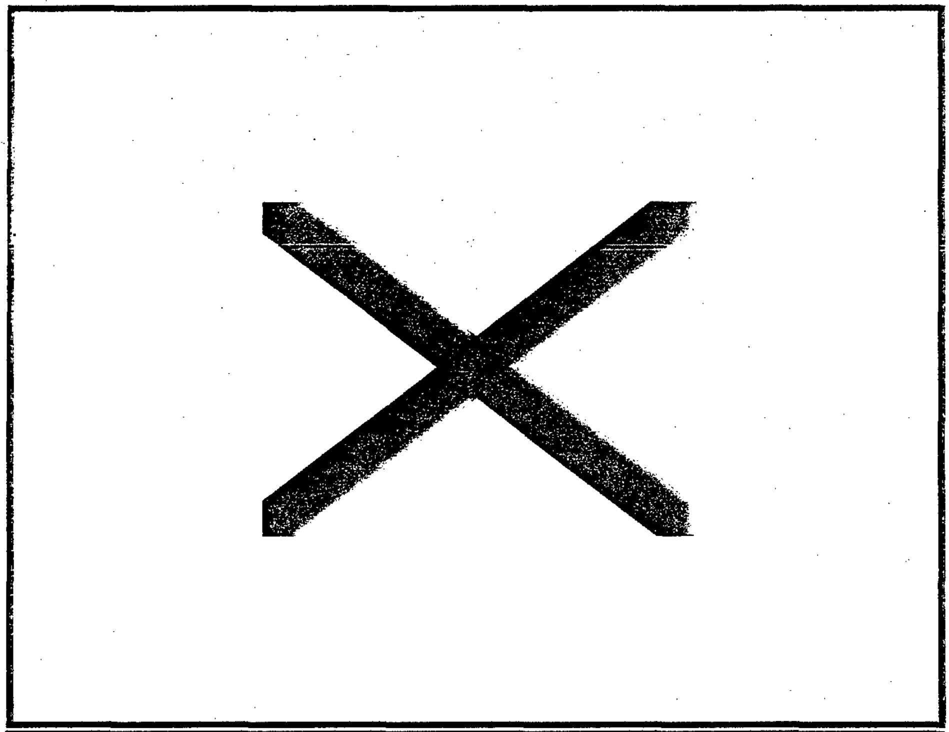


Figure 7-4.

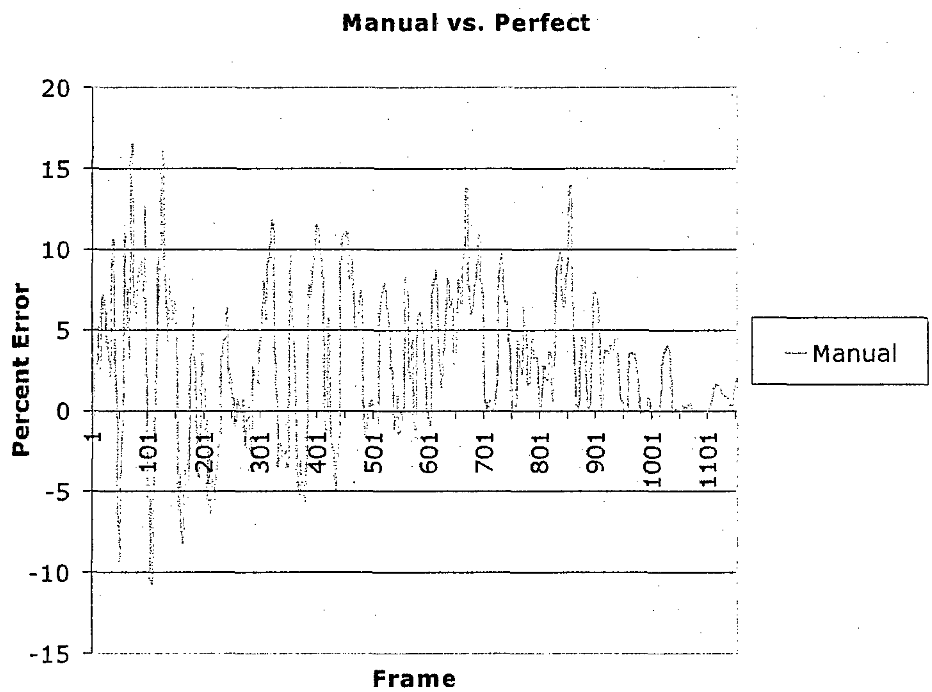
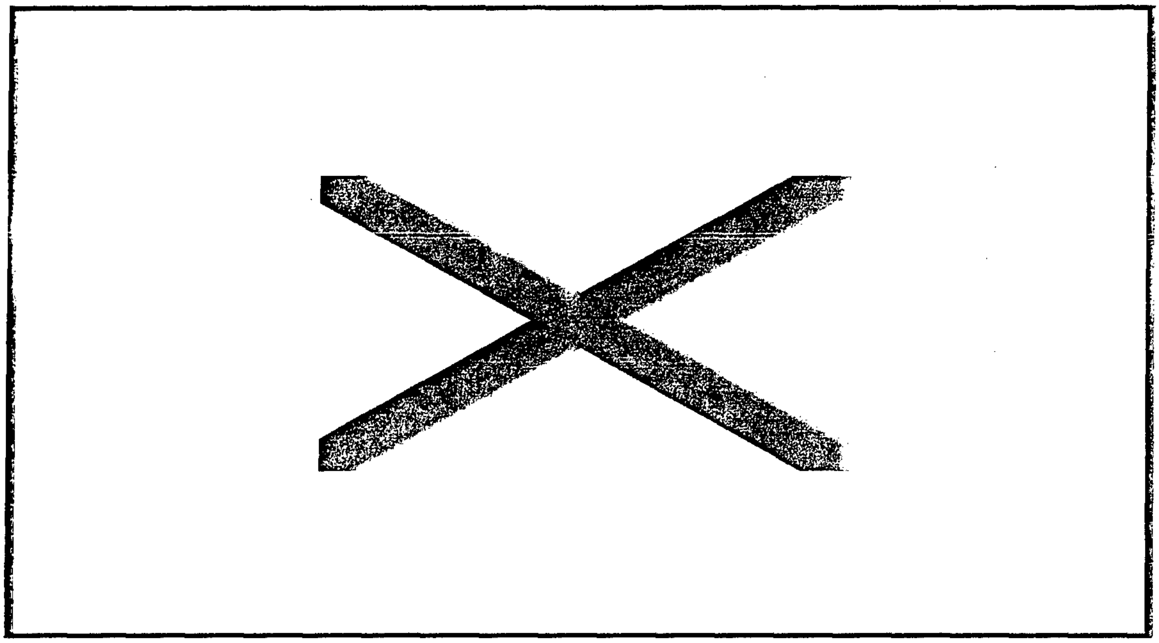


Figure 7-5.

An interesting point about Figure 7-5 is that most of the TM values generated from the manual ground-truth are within 10% of the values from the perfect ground-truth. In fact, the average error over all frames is 2.96%. Figure 7.6 shows the error and frequency of each key frame entry.



With these respectable results in hand we decided to increase the number of keyframes to see how accurate we could get the TM. In review it took approximately one hour for the initial keyframing exercise, resulting in 81 keyframes. By adding an additional 76 keyframes at a cost of 45 minutes of operator time, the average TM error over all frames was reduced to 2.45%. With the error dropping only 0.5% the added labor cost is hardly worth it.

8. SUMMARY

This paper describes an improvement of the GLCM TM that provides a more accurate prediction in the performance of modern terminal homing missile systems. This improvement consists of a normalized RSS composite metric value that not only measures the structural composition of the whole target compared to that of the surrounding background, as before, but also the most predominant feature on the target, such as a hot spot. The composite metric takes into account the potential for a tracker to track a single feature but with some level of penalization for only tracking a single feature of a multi-featured target. Additionally, the composite trackability metric can be calibrated for specific autotracker systems on the test range for Go/No-go test decisions.

The GLCM TM ground-truth accuracy study performed at MRDEC shows a dramatic change in TM values for a two-row shift or more in the ground-truth. It further shows that the difference between perfect ground-truth and a manual ground-truth aided with the new three dimensional target models in the TrackLab tool averages only three percent. This small loss in accuracy should be acceptable in most applications for the TM.

The TrackLab tools and the GLCM TM executable code are Government owned and available for Technology Transfer to private industry.

9. REFERENCES

1 B. Brackney, M. Helton, and R. Hammon, "Development of the Gray-Level Co-occurrence Matrix Target Trackability Metric for Imaging Infrared Missile Systems", *SPIE Proceedings, Image Processing*, Vol. 3377, 240-254, Orlando, FL. 1998.

2 G. Waldman, J. Wootton, G. Hobson, and K. Luetkemeyer, "A Normalized Clutter Measure for Images," *Computer Vision, Graphics, and Image Processing*, Vol. 42, 137-156, 1988.

3 W. K. Pratt, *Digital Image Processing, Second Edition*, John Wiley & Sons, Inc., 1991.

4 "Trackability Metric Improvement and Validation", Computer Sciences Corporation, N-TR-00-029, April 2000.

5 "Autotracker Design & Analysis", Computer Science Corporation, March 2002

6 "Design Modifications to the Gray-Level Co-occurrence Matrix (GLCM) Based Trackability Metric and the Resulting Performance", *SPIE Proceedings, Image Processing*, Vol. 3340, Orlando, FL, 2001.

Tracker Performance Metric

Teresa Olson, Harry Lee and Johnnie Sanders
Lockheed Martin Missile and Fire Controls

ABSTRACT

Currently, there is no standard quantitative measure of the performance of an image-based tracker algorithm. It is usually described as remaining in track for the entire run or losing track at some point during a run. This is purely a qualitative evaluation. Without a quantitative measure it becomes difficult to accurately evaluate the performance of a tracker algorithm. We have developed the Tracker Performance Metric (TPM) specifically for this purpose. It was designed to measure the output performance, on a frame-by-frame basis, using its output position and quality (sometimes referred to as confidence) state. The TPM can also be used as a measure of algorithm performance to compare against the Trackability Metric. The Trackability Metric was developed by AMCOM to determine how "trackable" a set of data should be. The TPM will be described and results presented.

Keywords: Tracker Performance Metric, Aimpoint Selection Metric, Quality Measure

INTRODUCTION

The purpose of this paper is to introduce a family of functions designed to yield standard quantitative calculations to rate performance for image-based tracker and aimpoint selection algorithms. These functions are adaptable to the algorithm evaluator's desired output requirements. They also have the ability to exploit composite tracker and aimpoint algorithm technology that may be used to leverage multiple algorithms within a solution suite. This approach requires a component track quality measure that represents the algorithm's inferred performance. We believe that a quality metric is an extremely important output of any tracker algorithm for producing robust solutions for today's military environment. An algorithm's ability to accurately assess its current quality state is a crucial metric for confidence measures that enable internal corrections and aid in external guidance decisions. If an evaluator does not have the ability to measure the quality of his tracker algorithm, the quality factor can be input as the maximum value of one. However, we must stress that this would not be considered an optimal solution. An accurate quality measure provides the most effective representation of tracker performance.

METRIC FORMULATION

At the center of this metric concept is the proposed function $f(x,y;Q,x_o,y_o)$, where (x,y) are image coordinate variables, Q is the algorithm's internally assessed quality metric (constrained by $0 < Q \leq 1$) and (x_o,y_o) is the algorithm's position output. The function is assumed to be integrable over all space, "centered" at the algorithm output position (x_o,y_o), and normalized to unit area ($\iint_{all\ space} f(x,y;Q,x_o,y_o) = 1$). The metric is calculated by integrating this function over the desired area of interest (A_D), typically centered on the true target location (x_T,y_T). In mathematical form, the performance metric is represented by $TPM = \iint_{A_D} f(x,y;Q,x_o,y_o)$. Because of the normalization criteria, the TPM is guaranteed to produce a value between zero and one ($0 \leq TPM \leq 1$).

The TPM was developed to be used in conjunction with the Trackability Metric (TM) or Gray Level Co-occurrence Matrix (GLCM). For detailed information on the TM, we refer the reader to [1,2,3]. The Trackability Metric measures how well a tracker should perform based on the tracker's frame-by-frame input imagery. On the other hand, the Tracker Performance Metric measures how well a tracker actually does perform on a frame-by-frame basis. Because both functions are normalized to the same scale (0-1), they are easily comparable. In general, a robust tracker would be expected to perform equal to or better than the GLCM representation.

To illustrate the utility of this proposed metric, it is simpler if we select a typical example of a set of desired output requirements for an algorithm. We will then be able to choose a representative for the function $f(x,y;Q,x_o,y_o)$ and the appropriate area of interest for integration. A typical performance goal for an image-based tracker is to reference all outputs to the centroid of the target under track. Large deviations of the tracker output from this centroid should be penalized as well as incorrect internal assessments of the tracker's quality measure. We have selected the area of interest as the entire target area but internal regions within the target itself are also viable candidates (especially for aimpoint selection evaluations). To review, we want to penalize tracker output deviations from the centroid of the target and also incorrect internal assessments produced from the tracker's quality metric evaluation.

In order to demonstrate this concept, choose the desired area of interest, A_D, to be a rectangular shaped target truth box having dimensions (L_x,L_y) that tightly encompass the target and is centered on the target centroid. Note that this region does not have to be rectangular. It could also be chosen to reflect the true target segment. As we will show later, choosing a rectangular region simplifies the computation. For our function, we choose a 2-dimensional multivariate Gaussian (normal) distribution, $N(\bar{r};\bar{\mu},\Sigma)$, centered at the tracker's output position $\bar{\mu}=\begin{bmatrix} x_o \\ y_o \end{bmatrix}$ with matrix $\Sigma=\begin{bmatrix} \left(\frac{L_x}{6Q}\right)^2 & 0 \\ 0 & \left(\frac{L_y}{6Q}\right)^2 \end{bmatrix}$, where $\bar{r}=\begin{bmatrix} x \\ y \end{bmatrix}$ and $0<Q\le 1$ is the confidence measure output of the tracker. The choice of the diagonal elements of Σ comes from our desire to have at least 99% of the area under the Gaussian surface contained within the target truth box when the Q value equals one ($\frac{L_x}{2Q}=3\sigma;\frac{L_y}{2Q}=3\sigma$). This gives the following functional form for $N(\bar{r};\bar{\mu},\Sigma)$:

$$
\begin{aligned}
N(\bar{r};\bar{\mu},\Sigma) &= \frac{1}{2\pi|\Sigma|^{0.5}}\exp\left(-\frac{1}{2}(\bar{r}-\bar{\mu})^T\Sigma^{-1}(\bar{r}-\bar{\mu})\right) \\
&= \frac{1}{2\pi\left(\frac{L_xL_y}{36Q^2}\right)}\exp\left(-\frac{(x-x_o)^2}{2\left(\frac{L_x}{6Q}\right)^2}-\frac{(y-y_o)^2}{2\left(\frac{L_y}{6Q}\right)^2}\right) \\
&= \frac{1}{\sqrt{2\pi\left(\frac{L_x}{6Q}\right)^2}}\exp\left(-\frac{1}{2}\frac{(x-x_o)^2}{\left(\frac{L_x}{6Q}\right)^2}\right)\cdot\frac{1}{\sqrt{2\pi\left(\frac{L_y}{6Q}\right)^2}}\exp\left(-\frac{1}{2}\frac{(y-y_o)^2}{\left(\frac{L_y}{6Q}\right)^2}\right) \\
&= N\left(x;x_o,\left(\frac{L_x}{6Q}\right)^2\right)\cdot N\left(y;y_o,\left(\frac{L_y}{6Q}\right)^2\right)
\end{aligned}
$$

The TPM is calculated by integrating this normal distribution only over the portion of the image that the target occupies. For a rectangular truthed region, A_D, this is represented by the following equation:

$$TPM = \iint_{target} N(\vec{r};\vec{\mu},\Sigma)\,dA = \iint_{target} N\left(x;x_o,\left(\frac{L_x}{6Q}\right)^2\right)\cdot N\left(y;y_o,\left(\frac{L_y}{6Q}\right)^2\right)dxdy$$

$$= \left(\int_{target_x} N\left(x;x_o,\left(\frac{L_x}{6Q}\right)^2\right)dx\right)\cdot\left(\int_{target_y} N\left(y;y_o,\left(\frac{L_y}{6Q}\right)^2\right)dy\right),$$

$$= \frac{1}{2}\left(erf\left(\frac{x_2-x_o}{\sqrt{2\left(\frac{L_x}{6Q}\right)^2}}\right)-erf\left(\frac{x_1-x_o}{\sqrt{2\left(\frac{L_x}{6Q}\right)^2}}\right)\right)\cdot\frac{1}{2}\left(erf\left(\frac{y_2-y_o}{\sqrt{2\left(\frac{L_y}{6Q}\right)^2}}\right)-erf\left(\frac{y_1-y_o}{\sqrt{2\left(\frac{L_y}{6Q}\right)^2}}\right)\right)$$

where $L_x = x_2 - x_1 > 0$, $L_y = y_2 - y_1 > 0$ and $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt$ is the standard error function used by Matlab (erf.m).

The crucial element in this formulation is that the function should be "centered" around the algorithm's output track point but the integration should only be taken over the desired area of interest. A mixture model of Gaussians matched to predefined lethality maps could be used to effectively quantify tracker performance in a more sophisticated aimpoint selection evaluation mode. As the output drifts away from the desired area or the tracker's quality metric becomes inaccurate, the area under $f(x,y;Q,x_o,y_o)$ will move away or expand into or out of region A_D.

EXAMPLES

Figure 1 illustrates TPM results for a tracker on US Army Aviation and Missile Command (AMCOM) data. The images show the truth bounding box and the 1-sigma, 2-sigma and 3-sigma ellipses of the normal distribution. In the left image, the TPM = 0.97 because the tracker's output is located near the center of the target and the tracker has a high confidence value. The right image has a TPM = 0.881, even though the output is located near the center of the target. This is because the tracker's quality (Q or TQ) is equal to 0.707. This has allowed more of the area to bleed outside of A_D. By design, a tracker producing an output on the target centroid, having a high degree of confidence (TQ = 1) will give a TPM approximately equal to one.

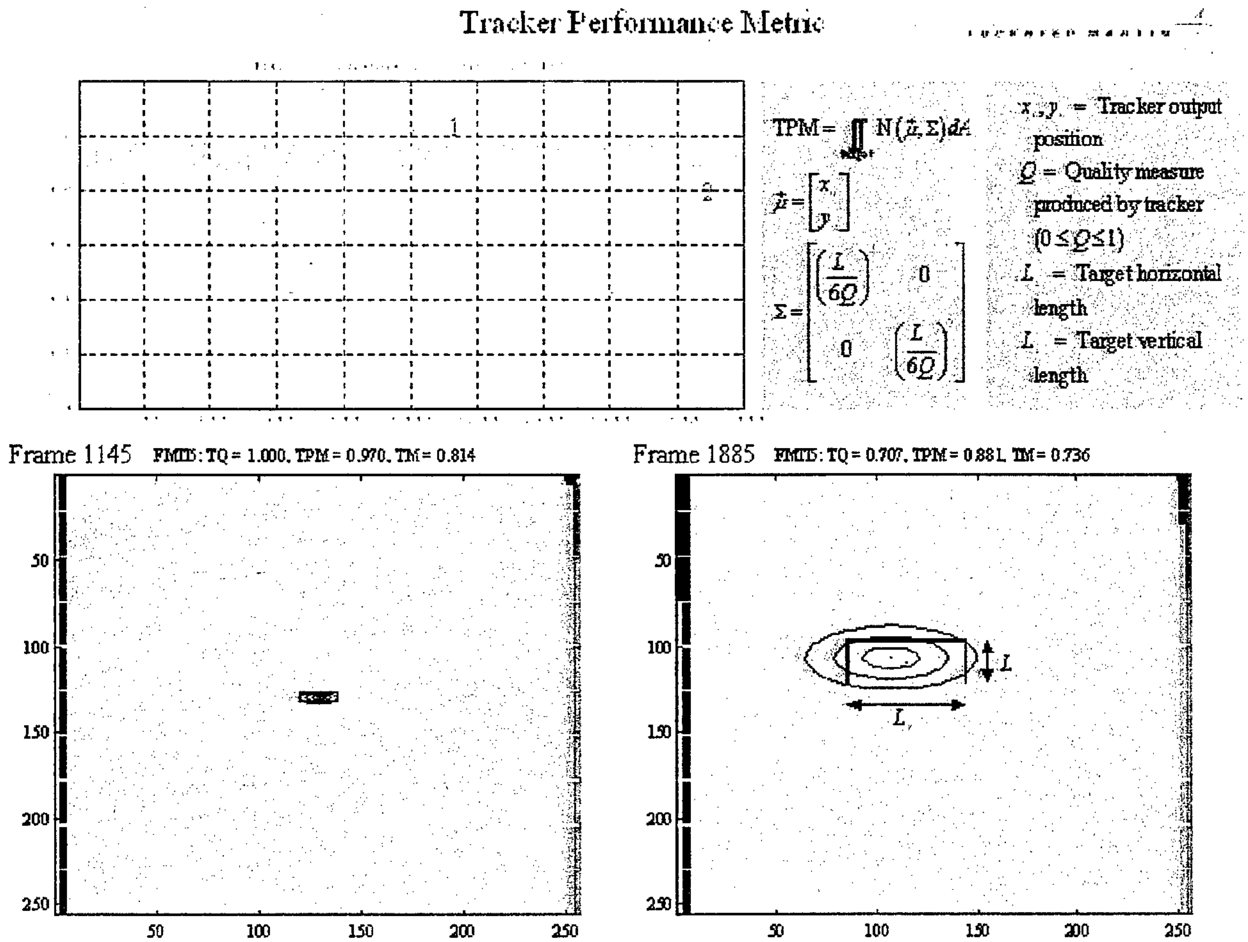


Figure 1: Tracker Performance Metric example 1.

Figure 2 illustrates TPM results for another tracker run on a different data set. Note the right image. The TPM has the very low value of 0.088. Not only has the tracker produced an output on this frame far from the target centroid, but also it is claiming a relatively high confidence of being on track (0.739). This is a problematic result for the tracker. The tracker is claiming good performance but actually producing poor results. Note that had the tracker been reliable enough to claim a poor quality rating, the ellipses would have expanded outward from the tracker output and towards the desired region of interest (the target). This would have yielded more area over the target region, and therefore, a higher TPM value. This result is fundamental to our metric formulation. Because the tracker has a better understanding of its performance limitations, it is more reliable and the algorithm performance metric reflects it. If the tracker can accurately identify when it has a lower confidence or quality, it can potentially indicate when to go into track or when not to update temporal information. A tracker that makes good assessments regarding its internal state, yields more accurate guidance decisions, and therefore, more accurate system results.

Figure 3 shows an example of the Tracker Performance Metric plotted against the Trackability Metric. It can be seen that the TPM corresponds well with the GLCM TM.

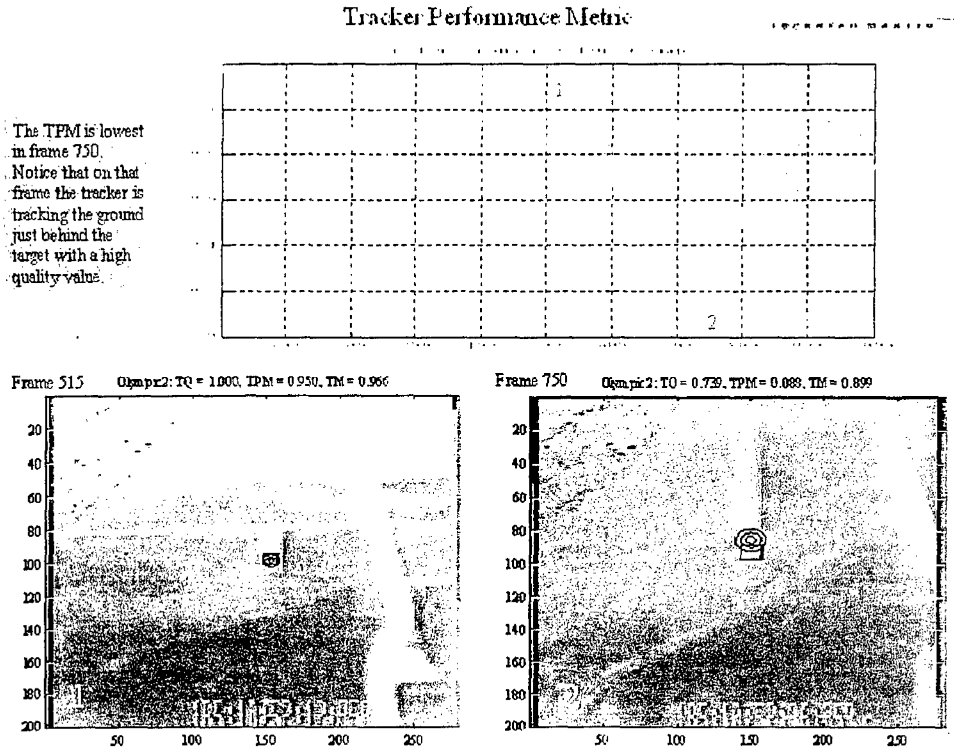


Figure 2: Tracker Performance Metric example 2.

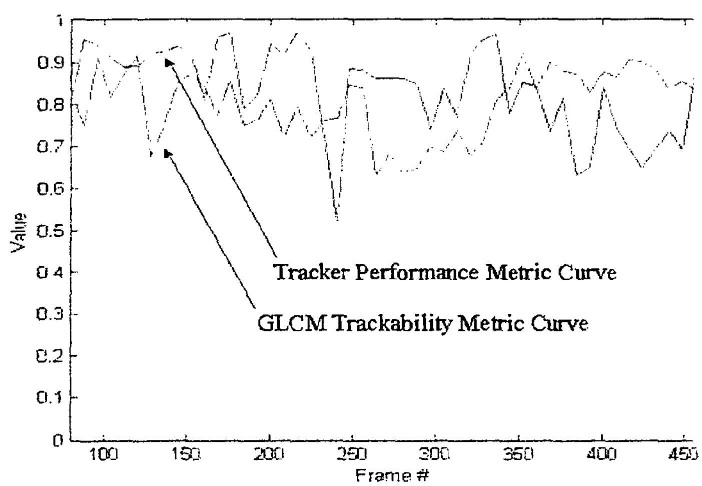


Figure 3: Tracker Performance versus Trackability Metric example.

SUMMARY

This paper has derived a quantitative method for evaluating the performance of any image-based tracker algorithm. This method, known as the Trackability Performance Metric, can be used to compare the performance of two or more independent tracker or aimpoint algorithms. It can also be used in conjunction with the Gray Level Co-occurrence Matrix Trackability Metric to determine how well a given algorithm performs verses how well it should perform on a given set of data.

ACKNOWLEDGEMENTS

The authors would like to thank Jamie Cannon, Kim Pham and Quang Tran for their dedicated work in truthing imagery and implementing the algorithms in Matlab. They would also like to thank Dr. Richard Sims from US Army Aviation and Missile Command (AMCOM), Redstone Arsenal, Alabama for providing the imagery being used in this metric development as well as Dr.'s Monte Helton and Ricky Hammon for providing the Trackability Metric code.

REFERENCES

1. Monte K. Helton, Ricky Hammon and Brian Brackney, "The development of the gray-level co-occurrence matrix target trackability metric for imaging missile systems", Proc. IRIS Passive Sensors, 1998, Vol. 1, pp. 195-207.
2. Brian A. Brackney, Monte K. Helton, and Ricky K. Hammon, "Development of the Gray-Level Co-occurrence Metric Target Trackability Metric for Imaging Infrared Missile Systems", SPIE Proceedings, Image Processing, Vol. 3377, 240-254, Orlando, FL, 1998.
3. Robert L. Hall, Ricky K. Hammon, Brian A. Brackney, and Benjamin J. Schmid, "Design Modifications to the Gray-Level Co-occurrence Matrix (GLCM) Based Trackability metric and the Resulting Performance", SPIE Proceedings, Image Processing, Vol.4365, 34-45, Orlando, FL, 2001.

METHODOLOGY FOR EVALUATING AUTONOMOUS IR TRACKER PERFORMANCE

2002 Ground Targets Modeling and Validation Conference
Houghton, MI, August 5-8, 2002

Paul D. Lavallee, Dr. Jay A. Lightfoot
Mark A. Chambliss, Carrie Kimbel, Judson R. Griffin III

Dynetics, Inc.
P.O. Box 5500
Huntsville, Alabama 35814-5500

ABSTRACT

Evaluating the capability of autonomous infrared (IR) trackers requires tools and methods to assess tracker performance as a function of target and background conditions. A methodology for assessing the performance of such IR trackers has been developed and implemented for analyzing captive flight test IR seeker imagery. Imagery is post-processed using the Tracker Analysis and Ground-Truth (TAG) tool to establish the target position in every frame of an IR sequence. The signature metrics of the target and background are extracted from each frame using the Software for Extracting Metrics from IR Sequences (SEMIRS). The sequences are then processed using the Imaging Seeker Algorithm Testbed (ISAT) to assess the performance of individual trackers. Finally, software has been implemented to match tracker performance results with the extracted signature information, producing a spreadsheet compatible file that can be used for analysis and plot generating. This process was developed and demonstrated using more than 50 IR sequences of a CFT IR seeker in closing encounters against ground vehicles.

1. Introduction

The evaluation of autonomous IR trackers has been ongoing for many years. Recent work at the US Army Aviation and Missile Command Research Engineering and Development Center (AMRDEC) has involved analyzing the performance of multiple IR trackers using Captive Flight Test (CFT) data. The desire has been to determine the correlation between tracker performance and target and background signature metrics. Dynetics and AMRDEC have derived a method for this type of analysis that is systematic and repeatable for any number of IR trackers or targets that need to be evaluated. Though developed for the application of evaluating CFT data, the methodology is applicable to evaluation of IR trackers using synthetic data.

Until recently, extracting the target and background signature metrics was a tedious, manually intensive process requiring many days to process a single sensor-to-target engagement. To overcome this, Dynetics developed the TAG tool to aid in ground-truthing IR image sequences (Reference 1). Ground-truthing is the process of identifying the target pixels in every frame of an IR sequence. The TAG tool allows ground-truthing entire sequences containing 4000 images in a matter of hours. The remainder of this paper will present the methodology and give some examples of its implementation.

2. Methodology for Evaluating IR Tracker Performance

The approach to evaluating IR tracker performance is shown in Figure 1. The process starts with the actual sequence of IR images taken from a CFT closing engagement of a seeker with a target. Also needed are the blackbody imagery and physical temperatures that are used for calibrating the imagery. Each of the individual steps of the methodology is addressed in the following subsections.

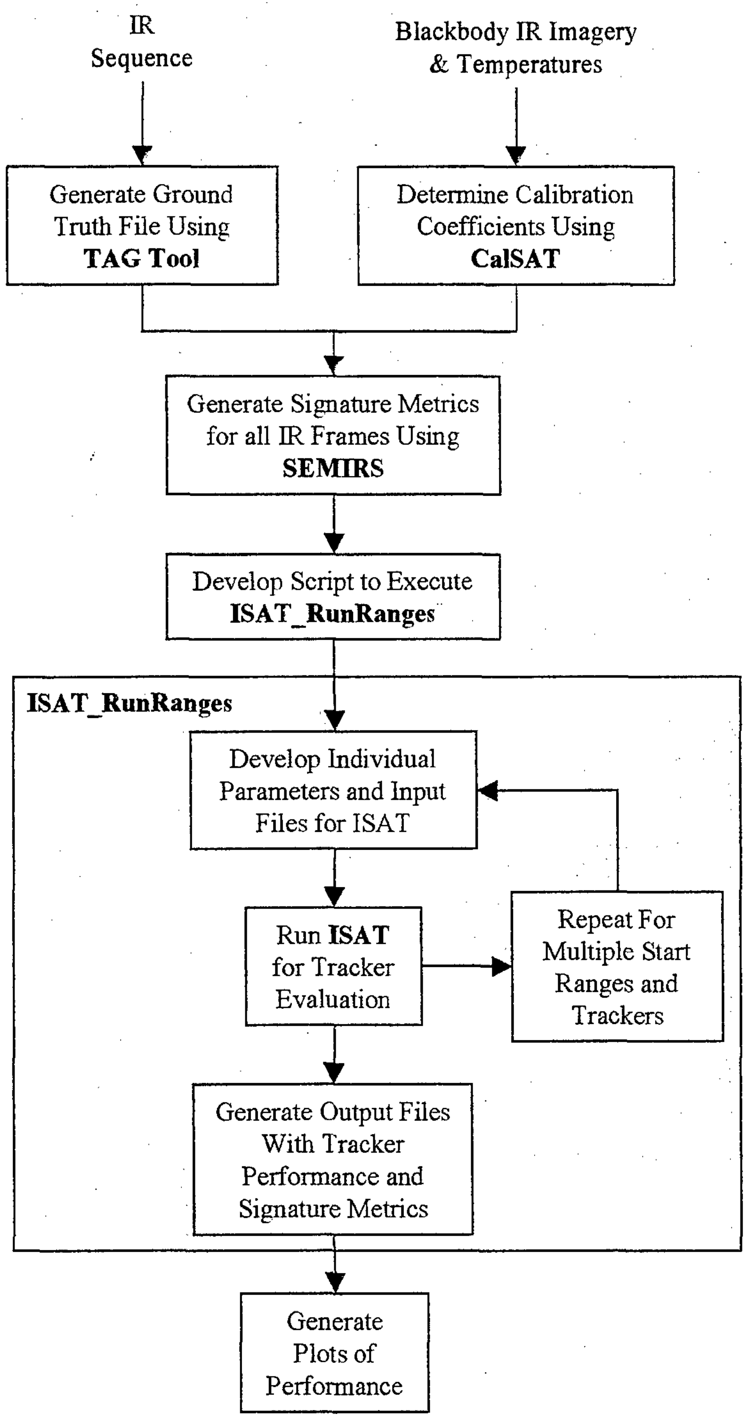


Figure 1. Methodology for Evaluating Autonomous IR Tracker Performance

2.1 Generating a Ground-Truth File

The process of generating a ground-truth file is an important step in evaluating an IR tracker. The ground-truth serves three purposes. The first is to target and background metrics in an automated fashion. The target pixels are identified within a rectangular box that a user defines using the TAG Tool. The second is to generate image stabilization information that can be used by the tracker in the event that there are large jerks from the gimbals in the CFT data. (The CFT data is typically collected open loop, and depending on the seeker gimbals, pointing joy stick, and seeker operator interactions, there can be inadvertent jerks in the seeker pointing.) The third purpose of the ground-truth is to enable automatic scoring of the IR tracker performance. This is accomplished by comparing the location of the IR tracker box to the ground-truth box.

Generation of the ground-truth for a single IR sequence comprising 4000 frames of data requires approximately 2-4 hours of time are required. The accuracy of the ground-truth box location and size is very important to the overall process including both the signature metrics extraction and IR tracker scoring. If the box is too big or is not centered on the target, then the target signature will be computed on a region that actually contains target and background. This same inaccuracy of the target box size or its placement will cause inaccuracies in the automatic scoring of the IR tracker. Figure 2 shows a ground-truth box that is appropriately sized and placed on the target.

To avoid these inaccuracies, great care is taken in generating the ground-truth using the TAG Tool. Often, the user will start ground-truthing the target location at the end of a sequence at which point the seeker was at its closest approach to the target. The user then typically steps backwards through the sequence, identifying the target in each image. As the target slant range increases and the target becomes smaller. the user is better able to precisely place and adjust the box over the target when performing this operation while stepping the IR images backwards. Reference 1 gives a complete overview of the TAG Tool and the steps for generating the ground-truth file.

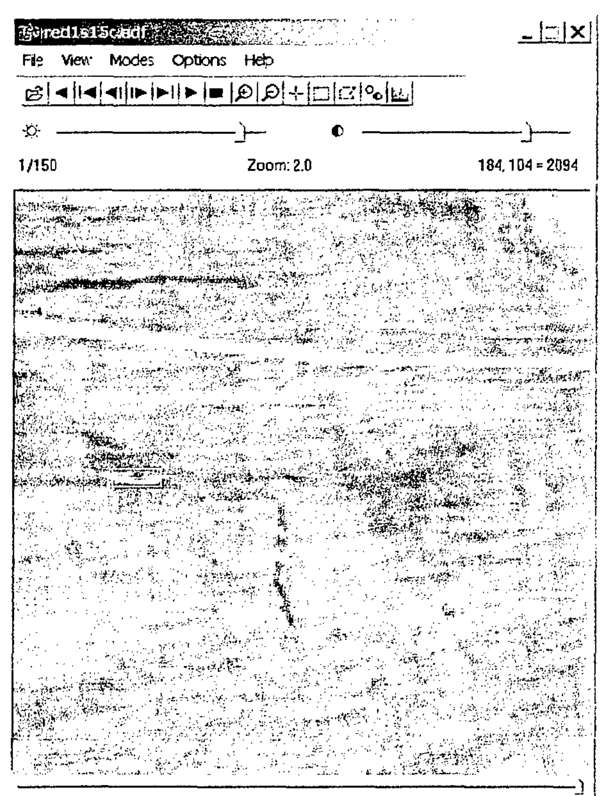


Figure 2. Example of Ground-Truth Box Placed on Target

2.2 Computing Calibration Coefficients

In order to compute signature metrics of the target and background, calibration information must first be determined. The typical process is to relate the IR image gray levels to temperature with a linear transformation. Over a small enough temperature range (20-30 K span), the linear fit is satisfactory for this transformation. The Calibrated Signature Analysis Tool (CalSAT) is used to generate a slope and intercept that transforms the individual pixels of an IR image into a temperature. Typically, the user imports an IR image with the four blackbodies for which the temperatures are known, as shown in Figure 3. The assumption is that the seeker viewed the blackbodies in the field during the CFT and close in time to the actual IR sequence to be analyzed. Also important is that the seeker digital imagery is stored without any automatic gain control adjustments to the data. Finally, a non-uniformity correction for a focal plane array seeker is necessary. CalSAT will compute the average gray level of each blackbody, and using the known temperatures of each, will compute a slope and intercept to transform the gray levels into temperature. The slope and intercept are stored in a file for later use.

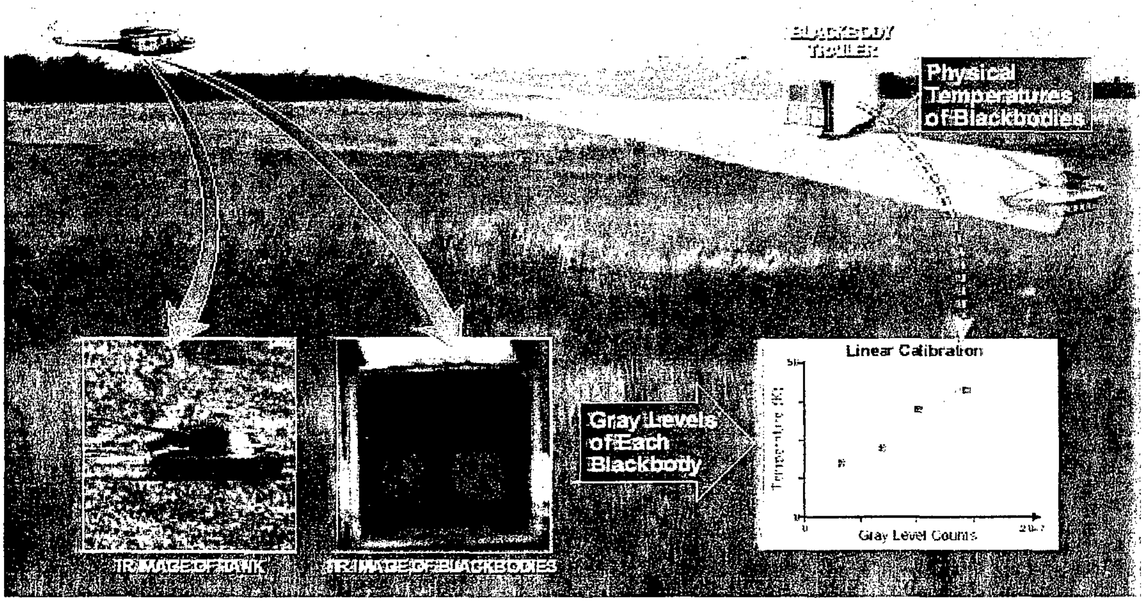


Figure 3. Test Range Scenario for Collecting Calibration Data

2.3 Generating Signature Metrics Using SEMIRS

The next step after generating a ground-truth file and calibration coefficients is to generate signature metrics for every image of the IR sequence. The Script for Extracting Metrics from IR Sequences (SEMIRS) was developed to accomplish this task. The inputs to SEMIRS are the IR sequence, its ground-truth file, and the calibration coefficients. SEMIRS then generates an output file containing many standard target and background signature metrics as shown in Table 1. In every case, the target region is defined as a rectangle (as discussed in the previous section) and the background region is an area that is twice as high and wide as the target area, centered on the target but excluding the target box. The metrics shown are apparent quantities, not source measurements, because this is typically what is of interest in assessing the IR tracker performance. SEMIRS is executed as a command line entry with associated arguments making it well suited for batch processing. Reference 2 gives a more in depth overview of the metrics shown in Table 1. Reference 3 gives more details on calibrating IR sequences and extracting signature metrics.

Table 1. Signature Metrics Computed by SEMIRS

Metric	Definition
$\overline{T}_{tgt}$	Average Target Temperature
σ_{tgt}	Standard Deviation of Target Temperature
N_{Tpix}	Number of Pixels in Target Region
M_{tgt}	Median of Target Temperature
$\overline{T}_{bkg}$	Average Background Temperature
σ_{bkg}	Standard Deviation of Background Temperature
N_{Bpix}	Number of Pixels in Background Region
M_{bkg}	Median of Background Temperature
ΔT	$\Delta T = \overline{T}_{tgt} - \overline{T}_{bkg}$
ΔT_{RSS}	$\Delta T_{RSS} = \sqrt{\Delta T^2 + \sigma_{tgt}^2}$
SCR_{RSS}	$SCR_{RSS} = \frac{\sqrt{\Delta T^2 + \sigma_{tgt}^2}}{\sigma_{bkg}}$
GLCM-TM	Gray Level Co-occurrence Matrix – Trackability Metrics

2.4 Running ISAT_RunRanges to Generate Inputs and Execute ISAT

The next step in evaluating the IR tracker is to execute a Perl script called ISAT_RunRanges. This script prepares the input files needed to run the IR tracker using the Imaging Seeker Algorithm Testbed (ISAT). ISAT is a software tool for design, development, and evaluation of acquisition and tracking algorithms developed previously by Dynetics for AMRDEC. The specific IR tracker under evaluation is added to and executed within the ISAT framework. An initial tracker box location is defined by the user after which the tracker autonomously determines where the target is on the next IR frame until the sequence is exhausted. For this specific application, the initial tracker box was taken from the ground-truth file. One newly added feature of ISAT is that it can automatically score the performance of the tracker by comparing the tracker box location to the ground-truth box. This is discussed in the next subsection. When a track is terminated either because the sequence was at its end or because the tracker lost track of the target, then ISAT writes out summary information. The ISAT_RunRanges script then uses this summary information to look up the corresponding signature metrics for the beginning and end of the run. It will also set up a new run at a closer slant range if the tracker failed to maintain track on the target throughout the entire run.

2.5 Automatic Scoring of the Tracker Performance

A metric was developed for assessing track quality or track error, and it is called "Instantaneous Track Error Metric – Overlap Error". It is based on the amount of overlap between the ground-truth and tracker boxes. Figure 4 shows a tracker box overlaid on a ground-truth box in an IR image with each region identified. The area (in pixels) for the truth region, tracker region, and overlap region are defined as NTRU, NTRK, and NOL, respectively. A ratio of NOL to NTRU gives an indication of the overlap compared to the ground-truth box. For the example in Figure 4, let us assume that this ratio is approximately 0.45, or there is 45% overlap between the tracker and truth boxes, as compared to the truth box. Another ratio between NOL to NTRK is defined to give some indication of how large the overlap region is compared to the track box, thus giving a measure of the size of the track box. For the example in Figure 4, let us assume that this ratio is 0.25, or there is 25% overlap between the tracker and truth boxes, as compared to the tracker box. If the track box is extremely large compared to the truth box, this second ratio will be much smaller than the first.

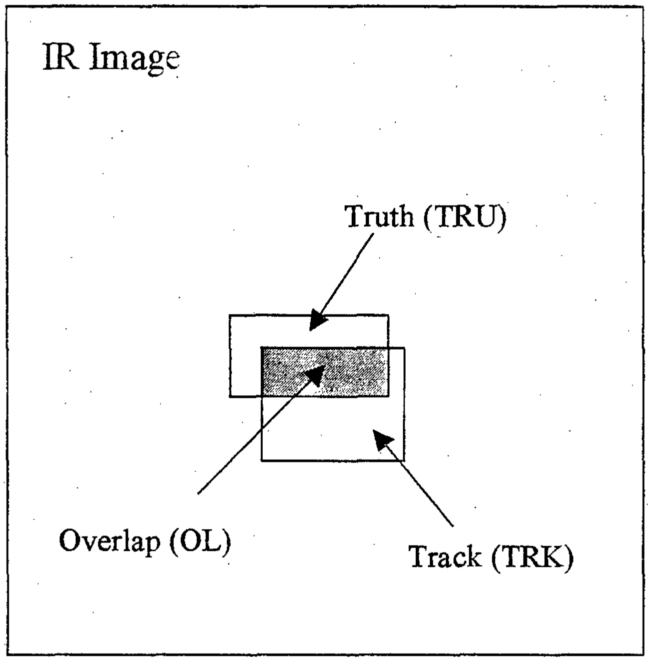


Figure 4. Track Box Overlaid on Ground-Truth Box in IR Image

Let us define each ratio as:

$$R_{TRU} = \frac{N_{OL}}{N_{TRU}}, \text{ and} \tag{1}$$

$$R_{TRK} = \frac{N_{OL}}{N_{TRK}} \tag{2}$$

By averaging the two ratios we can derive an instantaneous track quality metric as:

$$M_{TQ} = \frac{R_{TRU} + R_{TRK}}{2} \tag{3}$$

This metric has the desired characteristics that if the tracker box is exactly overlaid on and is the same size as the truth box, $M_{TQ} = 1$. If there is no overlap, the $M_{TQ} = 0$. To convert this metric into a "error" type metric where a "1" indicates maximum error and a "0" indicates no error, we can define an Overlap Error Metric as:

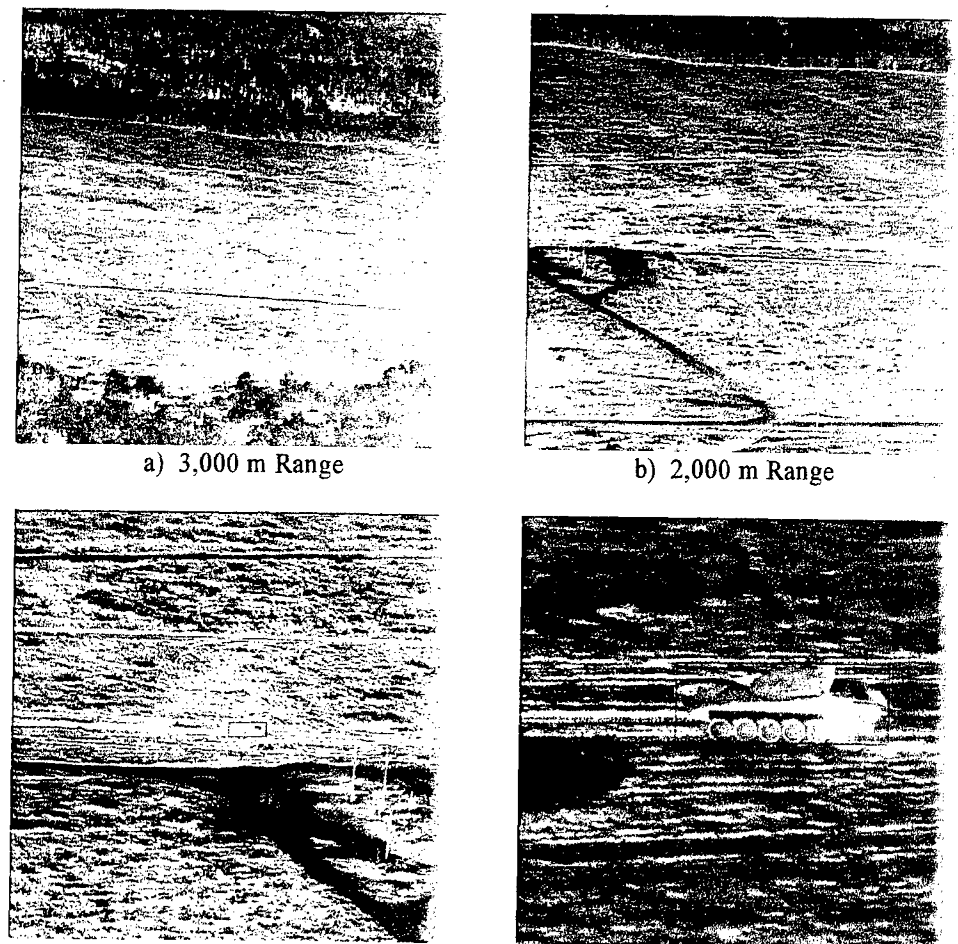
$$M_{OE} = 1 - M_{TQ} \tag{4}$$

As ISAT executes the tracker algorithm, it scores each individual IR frame using this overlap error metric. A criterion of M_{OE} of 0.95 or greater was chosen to indicate that the tracker was no longer tracking the target. If more that 10% of the total number of frames of the sequence exceeded this M_{OE} criterion, then the target was scored as "not tracked" by the tracker algorithm.

3. <u>Example of Methodology</u>

This section gives an example of using the methodology to evaluate an IR tracker performance against a T-72 tank. Imagery was collected from a captive flight test platform using an imaging IR seeker. The T-72 tank

was moving throughout the run and was being viewed by the seeker from a 270 degree target aspect. The tracker evaluated was a simple hot spot tracker. The initial track box is taken from the ground-truth data for this sequence, and the track was initiated at approximately 3,000 m slant range. Figure 5 shows representative images from the sequence at closing slant ranges with the ground-truth box overlaid on the target. The ground-truth box was generated using the TAG tool. Calibration was performed for this sequence using CalSAT along with a blackbody image and the known blackbody temperatures.



a) 3,000 m Range b) 2,000 m Range

c) 1,000 m Range d) 180 m Range

Figure 5. Representative IR Images of Closing Sequence at Various Slant Ranges

SEMIRS was then executed to extract the metrics for all of the IR frames of this sequence. Figure 6 shows a plot of ΔT_{RSS}, σ_{bkg}, and SCR_{RSS}. Note how the apparent ΔT_{RSS} increases as slant range decreases, as expected, because there is less atmospheric attenuation affecting the ΔT and there are more pixels on target which tends to increase σ_{tgt}. Figure 7 shows a plot of the target signature metrics including ΔT, σ_{tgt}, ΔT_{RSS}, and the number of pixels on target. The first three parameters are plotted using the left Y-axis, and the number of pixels is plotted using the logarithmic Y-axis on the right side of the plot.

Figure 8 shows a plot of the overlap error metrics (M_{OE}) as a function of range to the target. This demonstrates how well the hot spot tracker was able to keep the track box centered on the target. Given that the hottest portion of the target is the exhaust towards the rear of the tank, the tracker box was never completely centered up on the target, therefore there is always some overlap error. Figure 9 shows the tracker and ground-truth boxes for one IR frame of this sequence, demonstrating how the tracker centered its box on the hot exhaust, as expected.

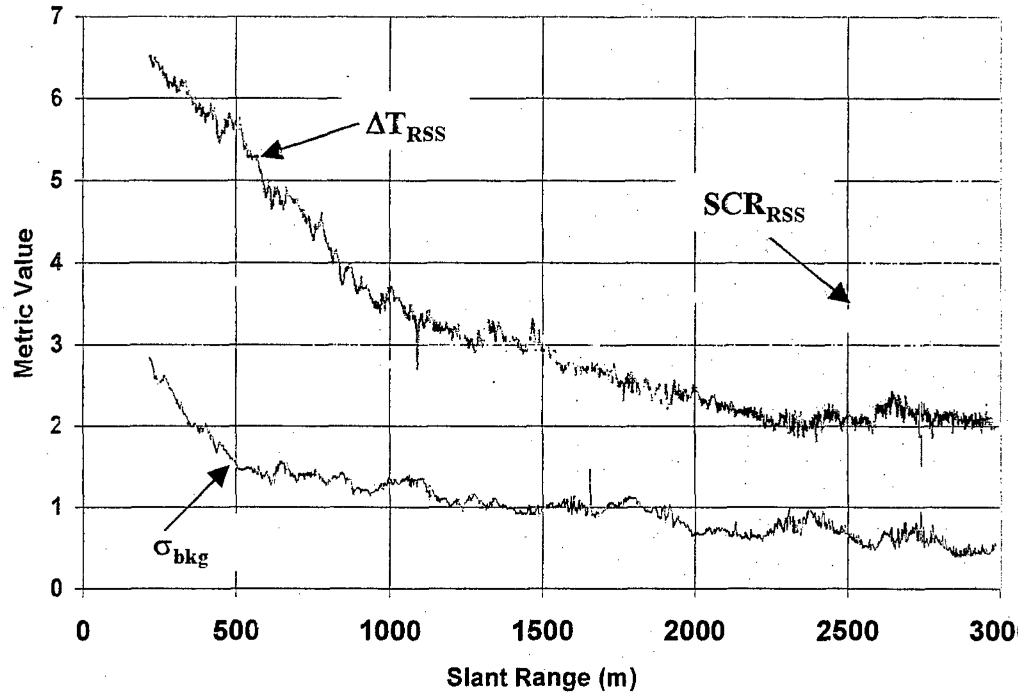


Figure 6. Contrast and Background Signature Metrics as a Function of Slant Range

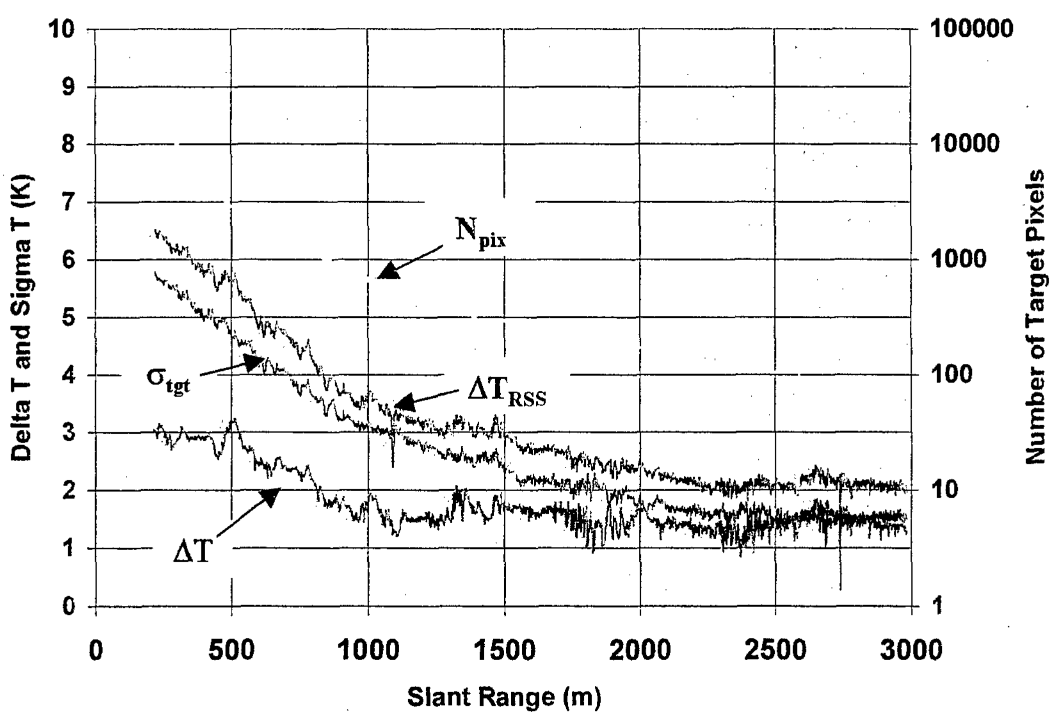


Figure 7. Target Signature Metrics as a Function of Slant Range

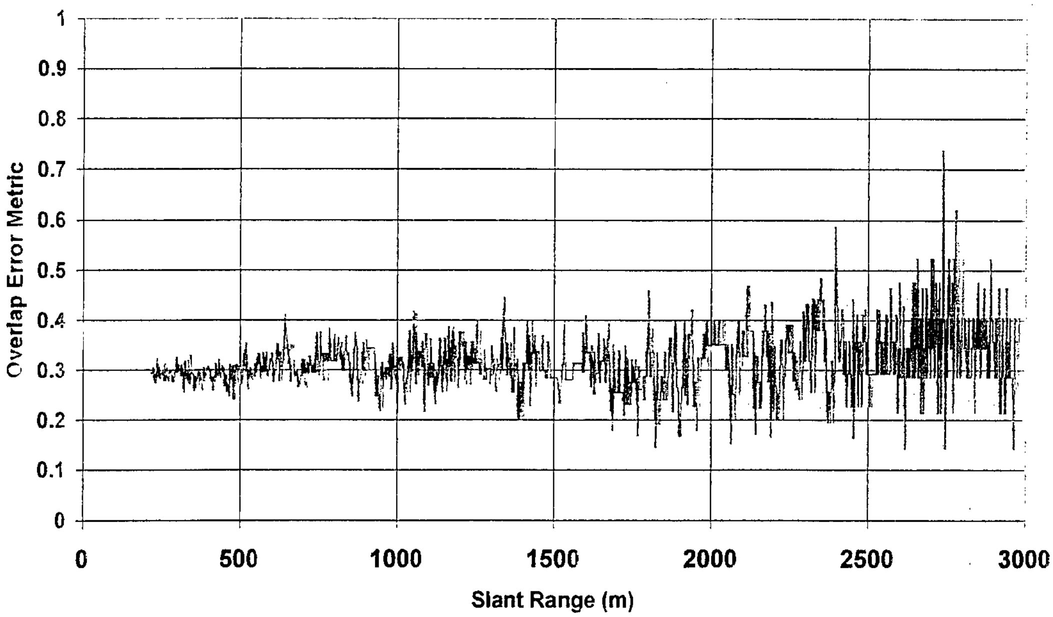


Figure 8. Overlap Error Metric (M_{OE}) as a Function of Slant Range

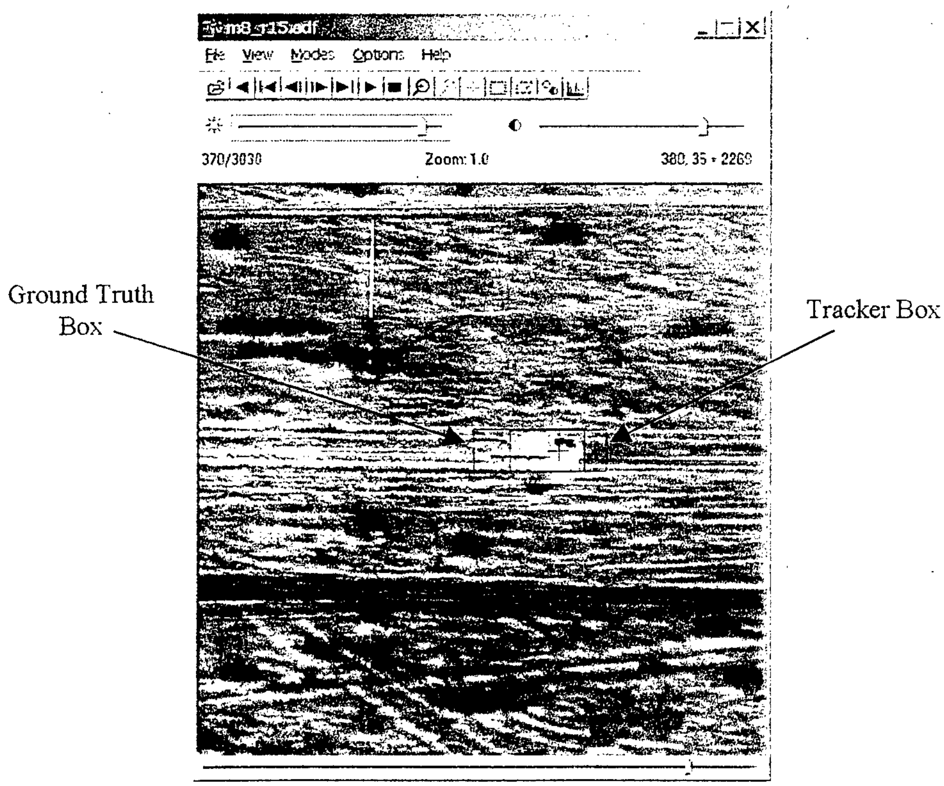


Figure 9. Tracker Box and Ground-Truth Box Placements

The ISAT_RunRange perl script accumulated the summary information shown in Table 2 for this particular sequence. The temperature metrics are in units of degrees Celsius. As a reminder, all values were computed as apparent quantities.

Table 2. Summary Signature Metrics at Beginning and End of Run

Metric	Beginning Value	Ending Value
$\overline{T}_{tgt}$	20.1601	30.8602
σ_{tgt}	1.5132	5.7718
N_{Tpix}	84	20436
$\overline{T}_{bkg}$	18.8259	27.8563
σ_{bkg}	0.4963	2.8030
N_{Bpix}	252	55486
ΔT	1.3342	3.0038
ΔT_{RSS}	2.0176	6.5067
SCR_{RSS}	4.0881	2.3213
GLCM-TM	0.8605	0.8375

By repeating this process for multiple seeker passes against an individual target, statistics of successful versus unsuccessful track attempts were accumulated. From this, analyses of signature metrics for successful tracks were compared to signature metrics for unsuccessful tracks.

4. Conclusions

The methodology presented in this paper was used to process over 50 different IR sequences collected during a CFT. The resulting data supported analysis of IR tracker performance as a function of target and background signature levels. Performance as a function of initial acquisition range was also derived from the data. The software tools were developed and enhanced to support the processing of a large volume of data in a relatively short period of time. The most labor-intensive portion of the processing is the ground-truth generation. Automation of the tracker processing and scoring was the key to completing the analysis on this large data set. The software tools developed by Dynetics for this particular type of analysis, including TAG tool, SEMIRS, CalSAT, ISAT, and ISAT_RunRanges, were crucial for automating as much of the process as possible.

5. Acknowledgements

This paper was based on work that Dynetics performed for the US Army Aviation and Missile Command Research Development, and Engineering Center (AMRDEC), Missile Guidance Directorate, at Redstone Arsenal, Alabama under the direction of Mr. Ron Passmore and Mrs. Lisa Cannon.

6. References

[1] Kimbel, Carrie, Jeff Todd, Paul D. Lavallee, Mark Chambliss, Jay Lightfoot, Jay Griffin, and Daniel Konkle, "Tracker Analysis and Ground-Truth Tool Description", Proceedings of the 2002 Ground Target Modeling & Validation Conference, 5-8 August 2002.

[2] Cannon, Lisa, Paul D. Lavallee, Mark Chambliss, and Eric Borg, "Signature Metrics for Imaging IR Seekers using Autonomous Tracking", Proceedings of the 12th Annual Ground Target Modeling & Validation (GTM&V) Conference, 7-9 August 2001.

[3] Lavallee, Paul D. and Mark Fowler, "Computing IR Signature Metrics Using Uncalibrated Field Test Imagery", Proceedings of the 12th Annual Ground Target Modeling & Validation (GTM&V) Conference, 7-9 August 2001.

Validation of Helicopter Flow Predictions Using Wind Tunnel LDA measurements

David Gatland, Dr. Lee Hayward, Dr Steve Rooks
Dstl, Air Systems,
A2 Bldg, 1064,
Farnborough, Hampshire
GU14 0LX

ABSTRACT

Helicopter engine exhaust plumes can constitute a large contribution to the total platform infra-red (IR) signature. The flow around such vehicles is particularly complex and an accurate representation of the plume is essential for high fidelity prediction of the IR signature. CFD techniques are well documented and have long been used for the generation of airflow and thermodynamic data. However. validation is essential for any problem to which the techniques are applied. Dstl uses the PHOENICS CFD code for the modelling of aircraft flow fields. In conjunction with PHOENICS, the SAPPHIRE modelling suite has been developed to aid model generation and IR signature analysis of air platforms.

This paper reports on the ability of SAPPHIRE to accurately predict the interaction of an exhaust plume with downwash caused by a helicopter rotor. Laser Doppler Anemometry (LDA) has been used to measure the flow around a simplified wind tunnel model orientated such that the free stream simulates the downwash. Comparison of this data with CFD results shows good general correlation and gives confidence that a 'fit for purpose' solution can be generated. Possible reasons for the discrepancies have been identified and are discussed.

INTRODUCTION

At Dstl Farnborough, the Air Systems Department includes a team concerned with the measurement, prediction and assessment of air-vehicle infra-red (IR) signatures. The SAPPHIRE suite of software has been developed to aid the generation and solution of flow models and IR signature predictions for various platforms. The on-going process of development and validation of SAPPHIRE has been reported regularly over recent years at the GTM&V conference [1,2,3,4].

The flow field relating to rotary wing aircraft is particularly complex and involves the interaction of the free stream flow, rotor downwash, tail rotor flow, engine intake flow and engine exhaust. In addition to these main features, the flow can also be affected by the proximity to the ground and the numerous inlets and outlets for items such as IR suppressors. Figure 1 shows the main features described above.

A key characteristic of the flow is the interaction between the rotor downwash and the engine exhaust, as this determines the position and size of the exhaust plume as well as potential hot flow impingement on the fuselage/tail boom. Accurate modelling of this interaction is important for any signature prediction and potential evaluation of IR exhaust suppressors. The SAPPHIRE helicopter models have been developed to account for such complex issues including the rotor flow field interaction. [1]

Validation of SAPPHIRE to model individual flow cases is an important and ongoing process. Well controlled wind tunnel tests are considered to be the main source for validation data. However, validation experiments for helicopters are few in number [5]. The main reason for this shortfall is the complexity inherent in measuring a 'real' helicopter flow field. Detailed flow field measurements are virtually impossible to make while a helicopter is in flight, while ground static measurements are rendered invalid due to the interaction of the rotor downwash with the ground. Therefore the process for validation of a code normally involves breaking down the problem into simple 'building block' experiments which look at the main factors individually.

In the mid 1990s, in support to helicopter infra-red suppressor assessments, a rotor downwash study jointly funded by UK MOD and Rolls Royce provided high fidelity flow field data for a generic helicopter/exhaust model. This study was specifically configured to look at the interaction of exhaust flows with rotor downwash.

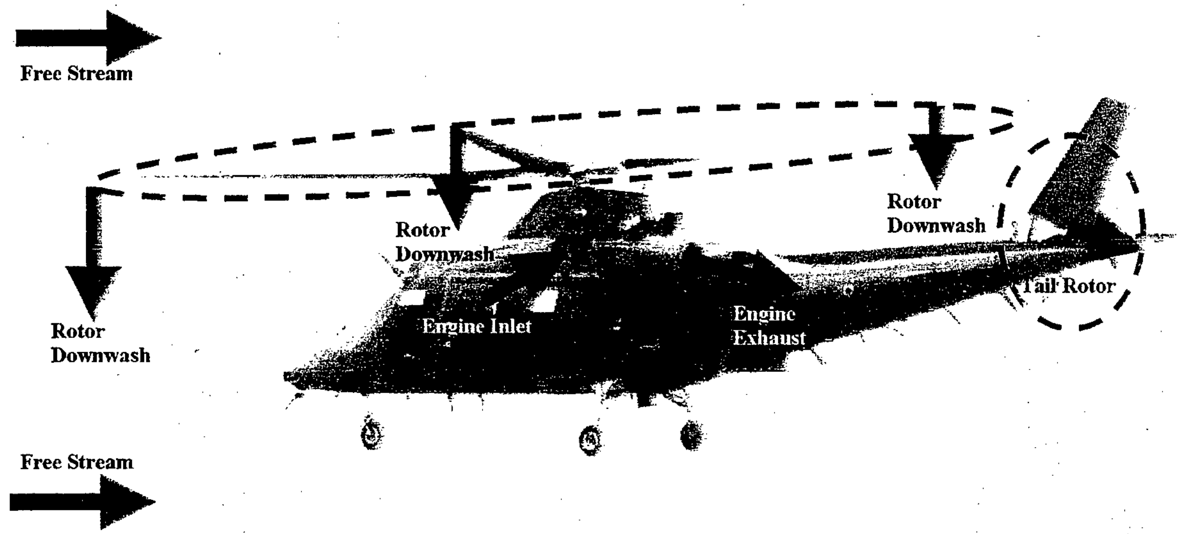


Figure 1: The Complex Flow Field Surrounding Helicopter Platforms.

The data from this study has recently been identified as a useful source for comparison with a SAPPHIRE helicopter flow field prediction. A CFD representation of the configuration has been produced and CFD predictions compared with the measured data. This paper reviews the numerical analysis and the related study carried out.

The aim of this study was to provide confidence that SAPPHIRE can predict the main flow characteristics resulting from the interaction of an exhaust flow with downwash. To be 'fit for purpose', the prediction needs as a minimum to show the general trends and position of the flow, along with the mixing and attenuation of the plume.

EXPERIMENTAL/COMPUTATIONAL METHODOLOGY AND TECHNIQUE

Wind Tunnel / LDA Setup

The Westland Helicopters Ltd (WHL) 12ft x 10ft low speed wind tunnel was used for the rotor downwash study. A generic helicopter model was mounted within the tunnel and orientated normal to the onset flow to simulate the rotor downwash. A separate flow was pumped through the model to simulate the engine exhaust. Figure 2 shows the exhaust and tail boom sections of the model mounted in the wind tunnel.

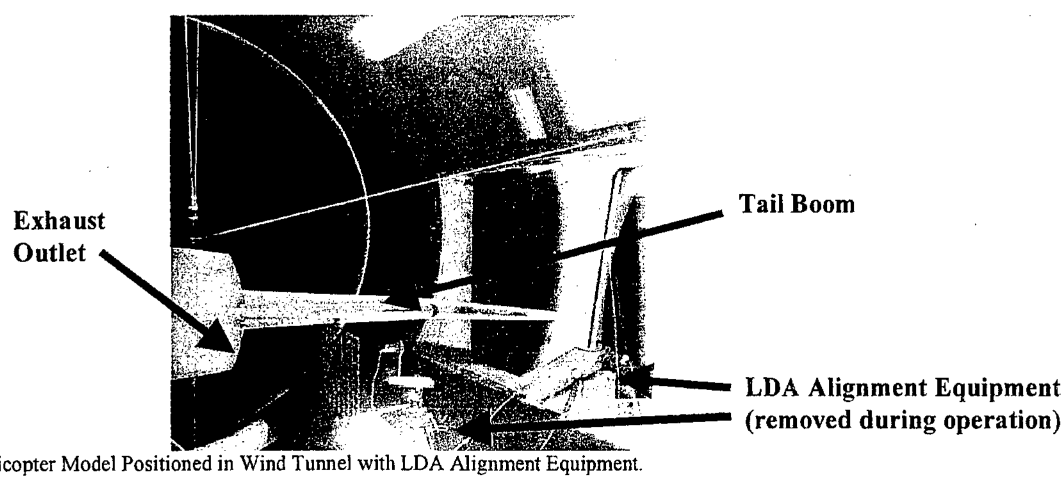


Figure 2: Helicopter Model Positioned in Wind Tunnel with LDA Alignment Equipment.

Laser Doppler Anemometry (LDA) was used to measure the flow velocity in four planes aft of the exhaust outlet. These planes were at the exhaust exit plane (y = 0)[1], 150 mm behind the outlet (y = 150), 300mm behind the outlet (y = 300) and 450mm behind the outlet (y = 450). A schematic is shown in figure 3. LDA requires the flow to be seeded with particles (typically smoke particles) and determines the velocity from the doppler shift undergone by the scattered light from six lasers. The main advantage of LDA over other methods of flow measurement (such as pitot/static probes) is the non-intrusive nature of the equipment with only the laser beams situated in the working section. This method does not require calibration and once the beams are aligned is highly resistant to experimental drift.

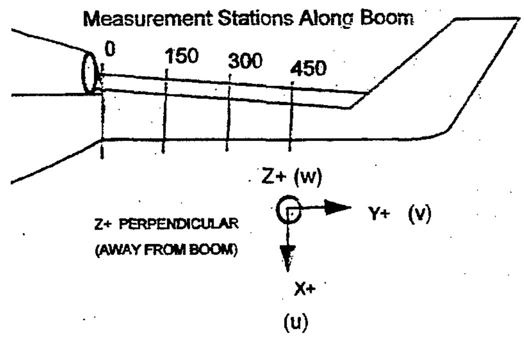


Figure 3: LDA Measurement Planes and Axis Definition.

SAPPHIRE/CFD model

a) General

For the CFD solution a mono-block fully structured grid was created from a CAD model using the FEMGEN module of SAPPHIRE as described in reference [1]. Generally, for problems such as this, an unstructured or multi-block CFD meshing technique and appropriate solver would be the preferred choice. Such techniques enable increased grid densities to be applied in areas of flow complexity or specific interest. Dstl has the ability to use such techniques with solvers such as FLUENT. However, at present these techniques have not been incorporated into the SAPPHIRE suite.

For a structured grid, memory constraints and computing power can be the major factor limiting the size and resolution of the grid. The CFD model is therefore confined to the port side of the geometry with a symmetry plane positioned along the vertical axis of the helicopter. The first stage of the investigation involved the creation of a CFD model with a relatively coarse grid. The aim of this was to generate a 'first pass' result that would indicate SAPPHIRE's general ability to model this case and highlight any potential problems without incurring the time penalty required for a higher resolution grid.

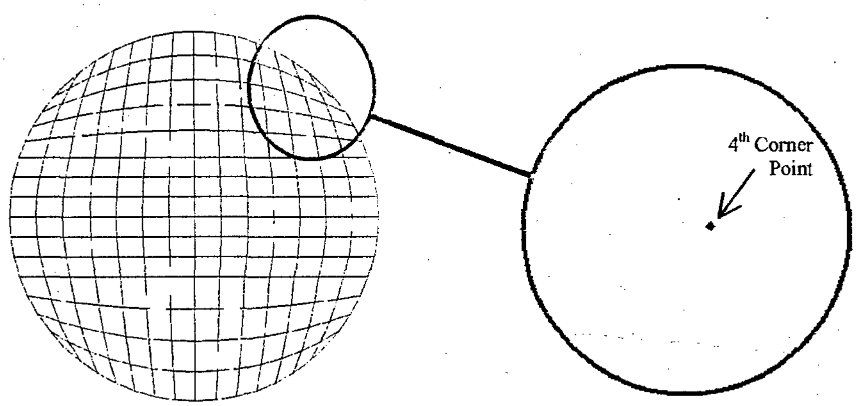
Once this stage had been completed, the geometry was adapted and the grid density increased significantly. Based on the coarse grid model, the flow around the forward section of the helicopter model was deemed to have no significant effect on the solution as the model represented an aircraft in hover. Therefore, in order to reduce the memory requirements and avoid cells with excessively high aspect ratio, the forward region of the original model was removed. The final grid (rear half) had 383104 computational cells compared to 66150 in the full preliminary model.

b) Issues

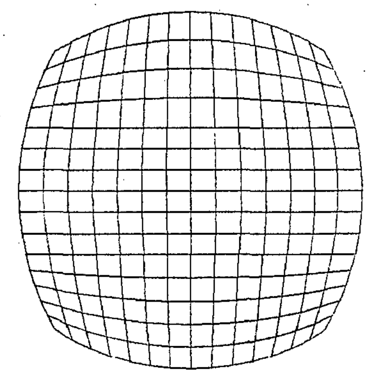
A number of turbulence models are available within the SAPPHIRE suite, including; fixed viscosity, k-W, standard K-ε, modified K-ε, Reynolds stress and renormalization group method (reference [6]). The standard K-ε turbulence model was used during the investigation as it has been validated for a wide range of flow cases and experience has shown it to give good results in this type of flow problem (reference [1]).

[1] The LDA data set at y = 0 is not the true exhaust exit plane but an XZ survey plane just aft of the nozzle exit. The true exit plane normal is angled away from the axial (Y) direction.

The 'body fitted' mesh is created by manipulating the 8 nodes of each hexahedral cell to give a good approximation to the original CAD geometry. Highly skewed cells and cells of high aspect ratio can cause the CFD solution to become unstable and diverge. Such cells are difficult to avoid completely due to the nature of a 'body fitted' mesh and the necessity of limiting the maximum resolution to the areas of interest only. As the grid density was increased, it was found that the turbulence model, which is less tolerant of these factors than the general flow equations, could not cope with the highly skewed cells generated at the 'corner' points of the circular exhaust geometry (figure 4a). For this case an approximation to the exhaust geometry was made as shown in figure 4b. The tail boom geometry also had to be modified in this way to enable the inclusion of the turbulence model with the fine grid.



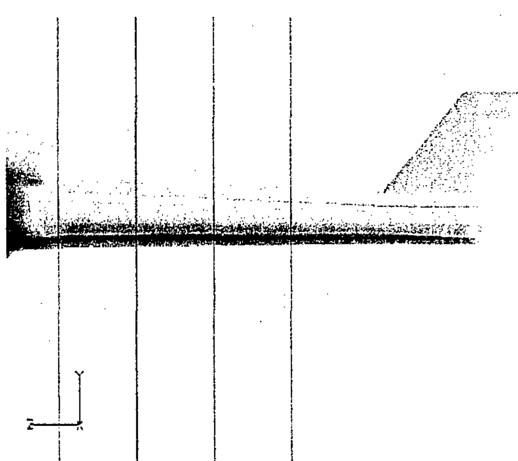
a) Round Geometry



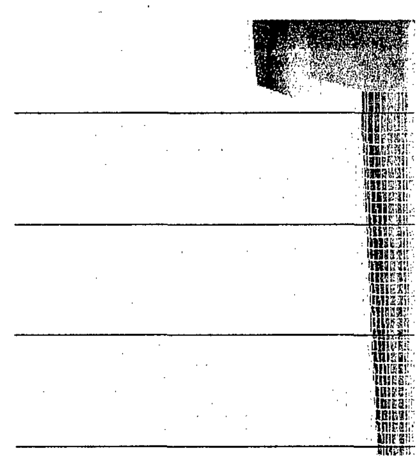
b) Modified Geometry

Figure 4: Exhaust Outlet Plane Exhibiting a Highly Skewed Cell.

Figures 5a and 5b show cross sections of the mesh and the measurement planes for comparison with the LDA data. The solid region is also shown. For this case, the only region of significant interest is that immediately surrounding, and aft of, the exhaust. The grid was created such that its resolution and quality was greatest in this region.



(a) XY Direction



(b) YZ Direction

Figure 5: Picture of CFD Solid Body, Grid Cross Section and LDA Measurement Planes.

RESULTS

Figure 6 shows the U, V and W Cartesian velocity components (see figure 3) given by the coarse CFD result, the LDA data and the fine CFD result at y=150. The 'squaring off' of the fuselage region in the fine grid model discussed above can be seen in figures 6c, 6f and 6i. The main features of the flow field at this point are marked on the LDA plot.

The flow in this plane exhibits a region of high downward velocity (region 1), probably due to the 'venturi effect' as the downwash is forced between the strong exhaust core and the tail boom. This feature can be seen in both CFD results. However, the definition is very poor with the coarse grid. The regions of greatest velocity are fairly similar for the fine grid result and the LDA data, with the peak region predicted further from the tail boom in the CFD model.

The slight recirculation of flow under the exhaust core region in the LDA result (region 2) is virtually unseen with the coarse model but is easily visible in the fine grid result. The accelerated region over, and to the left of the exhaust core (region 3) can be seen in both CFD predictions but is much more accurate in definition with the fine grid.

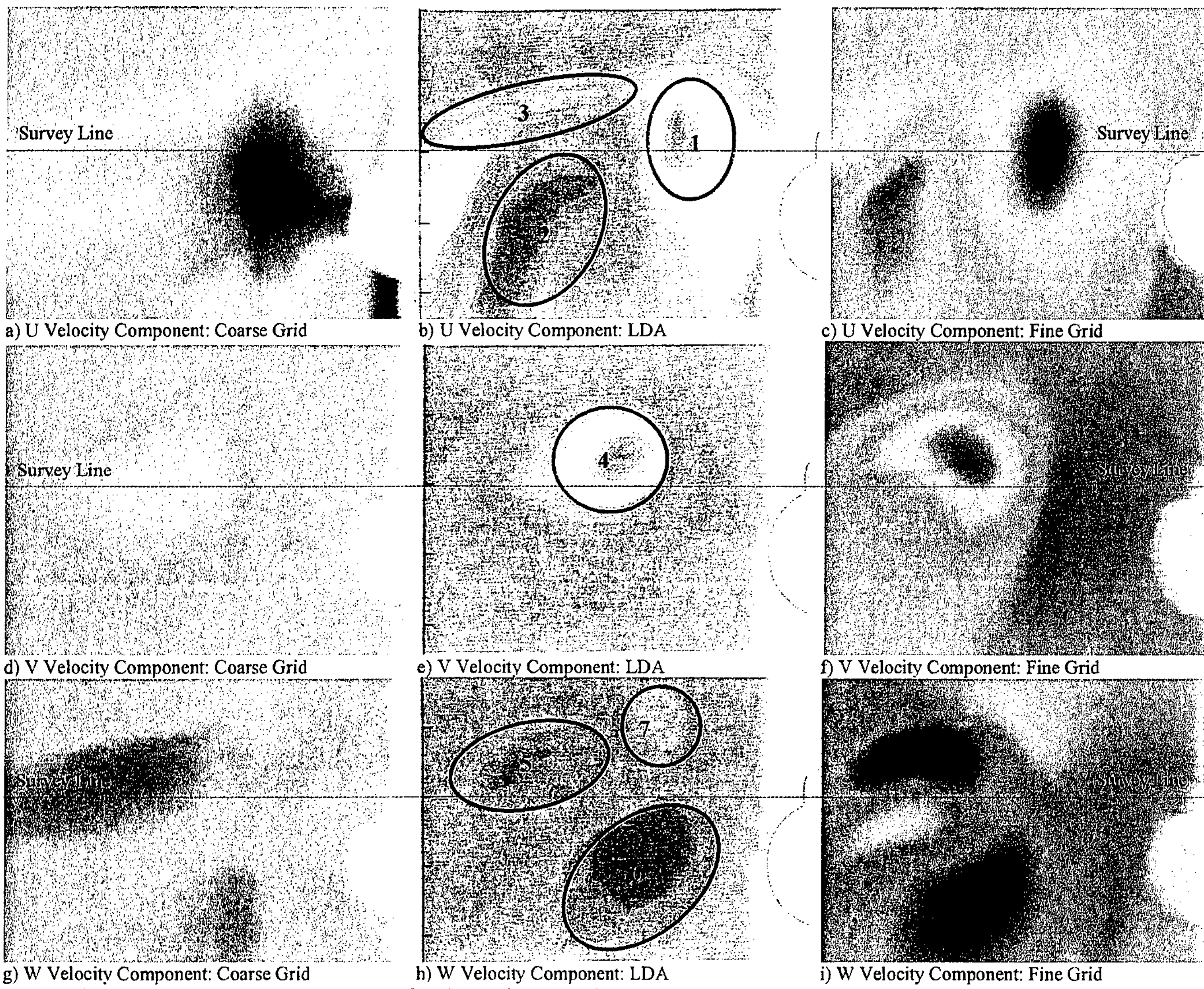
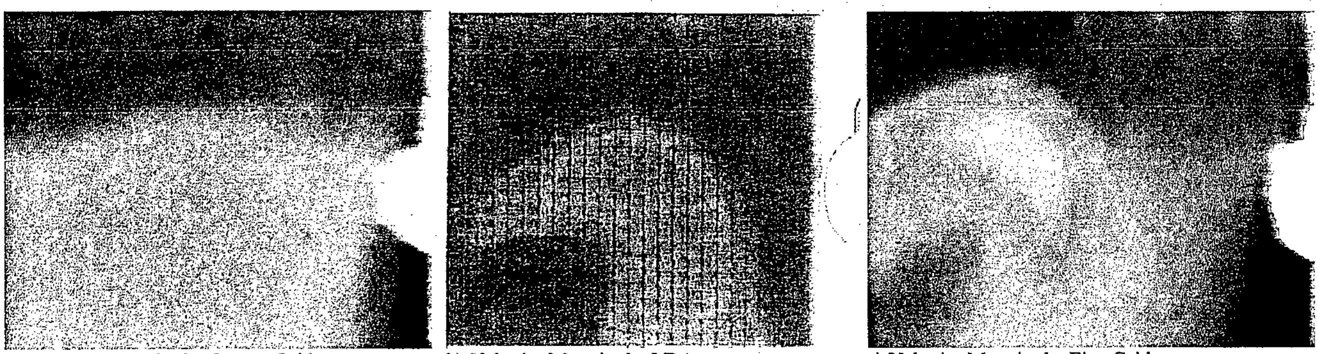


Figure 6: Coarse CFD Result, LDA Result and Fine CFD Result for U, V and W Velocity Components at y = 150.

The exhaust core region is clearly visible in figures 6d, 6e and 6f as the V velocity component is closely orientated with the exhaust direction. The fine CFD result shows a fair representation of the core region (region 4) with the predicted location offset further from the tail boom than in the LDA results. The coarse grid shows the core to be quite diffuse and indistinct.

The three main features exhibited by the LDA W velocity component (regions 5, 6, and 7) are predicted in both CFD models, again with the highest resolution given using the fine grid.

Figure 7 shows the total velocity magnitude given by the coarse CFD result, the fine CFD result and the LDA data at y=300. The preliminary CFD model shows a reasonable prediction of the exhaust flow velocity at this point but the shape of the exhaust flow is diffuse. The LDA data shows a distinct bowed region of flow that is duplicated in the fine CFD result. The fine result over estimates the exhaust velocity and predicts the plume slightly further out and higher relative to the tail boom than in the measurements.



a) Velocity Magnitude: Course Grid b) Velocity Magnitude: LDA c) Velocity Magnitude: Fine Grid

Figure 7: Course CFD Result, LDA Result and Current CFD Result for Total Velocity Magnitude at y = 300.

Figure 8 shows the velocity component and magnitude profiles along a survey line perpendicular to the tail boom in the plane y = 150 for each of the data sets, as marked in figure 6. Both the coarse and the fine grid result in a prediction that shows the general trends exhibited by the LDA data along the survey line.

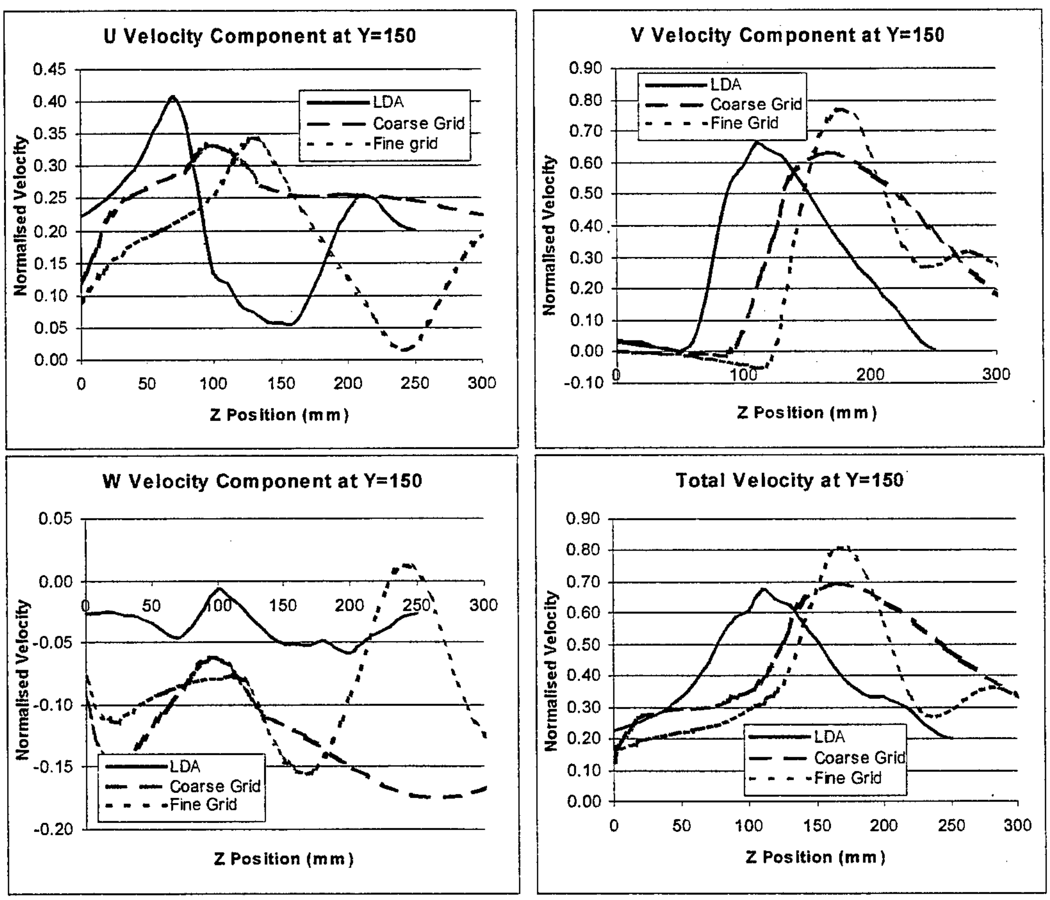


Figure 8: Velocity Components and Total Velocity Magnitude Along a Survey Line in the y = 150 Plane.

The velocity profile for the U component clearly shows the offset of the core exhaust region discussed above and the differences in the result using the coarse and the fine grid. Both CFD models predict the peak in downward velocity with the CFD showing this region to be further out relative to the boom than the measurements. The fine grid model follows the pattern exhibited by the LDA data quite closely while the coarse grid model fails to pick out the trough completely. This is partly due to the flow features being predicted lower relative to the survey line in the coarse grid result but the trough is still under estimated if the survey line is moved to compensate.

The V velocity component is roughly orientated with the exhaust flow and rises to a peak as the survey point passes through the exhaust core. Again the displacement outwards is seen in the CFD results. As the exhaust flow is roughly five times that of the downwash, the V component is dominant and hence the total velocity magnitude exhibits very similar trends and magnitudes as the V velocity component. The core velocity is predicted to be greater at this point than the measured result, showing the predicted plume is mixing and dissipating slower.

The W velocity component shows the worst correlation between predicted and measured data. Again the fine grid captures the peaks and troughs along the survey line better than the model using the coarse grid but both CFD results are greater in magnitude than the LDA data.

Figure 9 shows the U and V flow components along the axial centreline of the exhaust. The measured and predicted U and V velocity components show similar characteristics in both general trend and magnitude, though the CFD results generally lead the LDA line slightly. As has been discussed above, the predicted exhaust is flowing at a greater effective angle to the tail boom than the LDA data indicates. This means that the axial survey line passes out of the core exhaust region more quickly for the CFD results and hence the velocity components tend towards the free stream more rapidly.

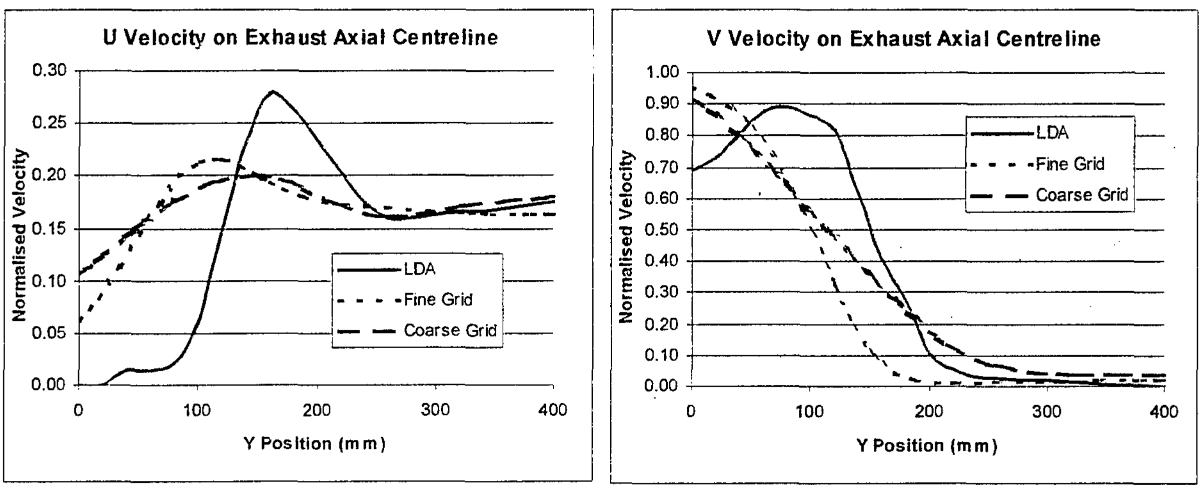


Figure 9: Velocity Components Along an Axial Survey Line.

DISCUSSION/CONCLUSION

This investigation has provided further confidence that SAPPHIRE is capable of predicting the flow field characteristics necessary to provide a fit for purpose thermal evaluation for helicopter signature analysis. While it has been shown that SAPPHIRE can predict the general flow features relating to the exhaust/downwash interaction, there are discrepancies between predicted and measured plume location and flow mixing. The solution is only as good as the boundary conditions used to define and constrain the flow field. A more accurate prediction of this case would require resolving several key features that are not available within the LDA data.

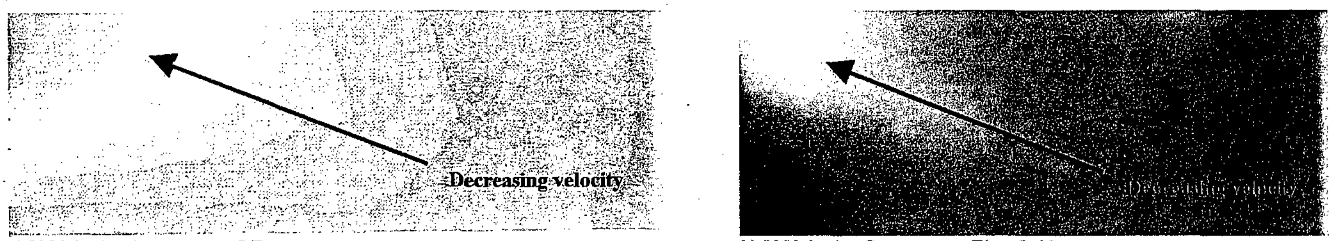
The most obvious factor is the boundary condition used for the exhaust in the CFD model. The exhaust boundary condition was set such that the outlet had a 'top hat' mass flow and a velocity vector perpendicular to the exit plane, i.e. the exhaust flow was at a constant magnitude and direction across the entire exit plane (figure 10a). In reality this is not going to be the case. The exhaust flow is piped through the model, undergoing several changes in direction and as such is unlikely to be

flowing perpendicular to the exit plane. Coupled with this, boundary layer effects within the exhaust duct will cause a velocity profile as shown in figure 10b.

a) 'Top Hat' Velocity Boundary   b) Profiled Velocity Boundary

Figure 10: Difference between 'Top Hat' and 'Profiled' Velocity Boundaries.

The CFD calculation domain was bounded by free boundaries while the wind tunnel used a closed section design. The helicopter model caused a significant blockage to the free stream flow giving a varied flow profile above the model. The CFD inlet boundary above the helicopter was again given a top hat value based on the wind tunnel velocity. The U velocity component plots at the effective rotor plane for the LDA and the fine grid model are shown in figure 11. This clearly shows the reduced downwash in both the experimental and computational results in this region. However, the LDA data shows a significantly higher decrease. The reduction in free stream flow would lessen the effect of the 'downwash' on the exhaust, particularly close to the fuselage and exhaust exit. This, together with the uncertainty of the flow conditions at the nozzle exit, is the probable cause of the discrepancy in exhaust flow position.



a) U Velocity Component: LDA   b) U Velocity Component: Fine Grid

Figure 11: U Velocity Component in Rotor Plane.

The slower mixing and attenuation of the exhaust in the predicted results than the LDA data may also indicate that the turbulence boundary conditions and predicted intensity used in the CFD calculations were lower than they were in reality. The use of a different turbulence model may also improve the correlation between the results and further work is being undertaken to investigate this factor.

Finally the position of the survey lines and planes for the CFD data set may not correspond precisely with those used for the LDA measurements. The reference point for the LDA measurements was a point just aft of the exhaust and is not precisely known.

ACKNOWLEDGMENTS

This work has been funded under UK MOD Applied Research Package 12 and carried out by Dstl Farnborough.

The 'Rotor Downwash Study', through which the wind tunnel measurements were generated, was jointly funded by UK MOD and Rolls Royce Plc. The Authors would like to thank Rolls Royce for the use of this experimental data (copyright Rolls-Royce Plc 1998).

The Authors would like to thank S&C Thermofluids for their support and help with SAPPHIRE.

REFERENCES

[1] Rooks, Steve. Smith, Anthony G., 1997, "Flow Field and IR Signature Modelling for Ground Targets", *Proceedings of the Eighth Annual Ground Target Modelling and Validation Conference,* Houghton, MI.

[2] Lavery, N. Brown, S. Spittle, J. Rooks. S. Hayward, L. Smith, A. 2000, "A Coupled CFD-Thermal Solver Routine for Detailed Platform IR Prediction", *Proceedings of the Eleventh Annual Ground Target Modelling and Validation Conference,* Houghton, MI.

[3] Lavery, N. Brown. S. Spittle, J. Rooks, S. Hayward, L. Smith, A. 2001, "Improved Modelling Capabilities of PHARO – A Coupled CFD-Thermal Solver for Detailed Platform IR Prediction", *Proceedings of the Twelfth Annual Ground Target Modelling and Validation Conference,* Houghton, MI.

[4] Newman, S. Rooks, S. Bull, C. Green, J. Taylor, K. Smith. A. 2001, "Validation of SAPPHIRE Heat Transfer Predictions", *Proceedings of the Twelfth Annual Ground Target Modelling and Validation Conference,* Houghton, MI.

[5] Research and Technology Organisation (RTO), 1999, "A feasibility Study of Collaborative Multi-Facility Windtunnel Testing for CFD Validation" RTO TR-27.

[6] Launder, B.E. Spalding, D. B. 1972, "Mathematical Models of Turbulence" *Academic Press.*

Application of a Monte Carlo Technique to Model Helicopter Plume Reflections

B. Jane Goldsworthy, S. Rooks and M. Fair
Dstl Farnborough, Hampshire
GU14 0LX, UK

D. Filbee
Insys Ltd, Reddings Wood
Ampthill, Bedford
MK45 2HD, UK

ABSTRACT

High fidelity infrared signature modelling of targets requires the accurate representation of not only factors concerned with the target itself but also the immediate environment. Modelling of multiple reflections from extended thermal sources can often be oversimplified, partly due to the complexity of the problem and the resulting intensive processing requirements. The use of Monte Carlo ray tracing techniques for reflection modelling is a documented approach. Insys Ltd has developed a code, McCavity, on behalf of Dstl based on these techniques, to quantify various aspects of the multiple reflections problem to the infrared signature.

This paper reports on the use of McCavity for the assessment of reflections arising from an exhaust plume onto a representative helicopter tail-boom and static rotor structure. A simple, generic, example is used to show that the contribution of reflections from an extended source to the infra red signature can be significant. McCavity code validation against standard published results is also presented.

INTRODUCTION

The Air Vehicle Signatures Group at Dstl Farnborough undertakes and sponsors for MOD, research into the measurement, prediction and assessment of military air-vehicle infra red (IR) signatures. In simple terms, there are two main contributory sources to the IR signature of an aircraft; the first comes from the plume and is therefore a direct consequence of the powerplant system, and the second comes from the airframe itself. The signature arising from the airframe is primarily attributed to thermal emission and surface reflections. The surface can become heated through several mechanisms, including aerodynamic heating, conduction from internal hot sources (electronics, powerplant), plume impingement and solar absorption. Plume impingement is a particular problem for helicopters, the effect of which can be seen in Figure 1, where the airframe structure is seen to be heated by the plume.

High fidelity IR prediction models are dependent on the accurate representation of the thermal and radiation processes. This requires appropriate definition of the target geometry, material properties, heat sources and the influence of the surrounding environment. Accurate representation of surface temperatures is particularly important and a vital element of the prediction process. At Dstl advanced CFD solvers and thermal modelling tools are employed to provide the thermodynamic data for aircraft exhaust gases and surface heating. These include SAPPHIRE [1], MuSES, FLUENT and PHARO [2].

The platform thermal characteristics determined by such models can be imported into radiation models, which can also determine the environmental influences on the platform. One such code, CAMEO-SIM, has an advanced synthetic scene generation capability, and has been the topic of previous GTMV papers [3, 4, 5]. CAMEO-SIM is used to generate radiometrically accurate images based on thermally embedded targets within an environment and is used to provide an accurate assessment of EO camouflage systems. However, CAMEO-SIM is currently limited to solid targets; consequently the associated plume is not calculated as a source and therefore reflections from this source are not modelled. For helicopter signature prediction this can be a particular limitation in the medium waveband. For example, Figure 2 shows an IR image in which a plume reflection can be seen on the main rotor blade. For such platforms, reflections arising from the high temperature and area of the powerplant have the potential to contribute significantly to the overall platform signature.

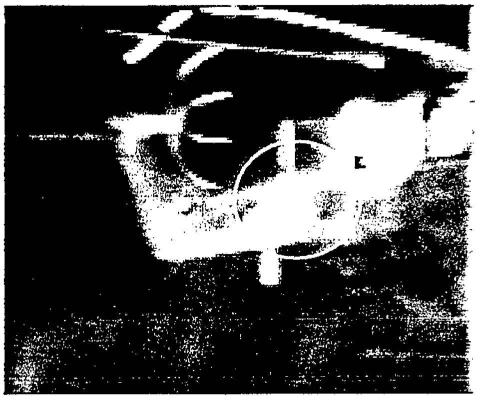


Figure 1 Measurement IR image showing tail-boom heating due to impingement of plume

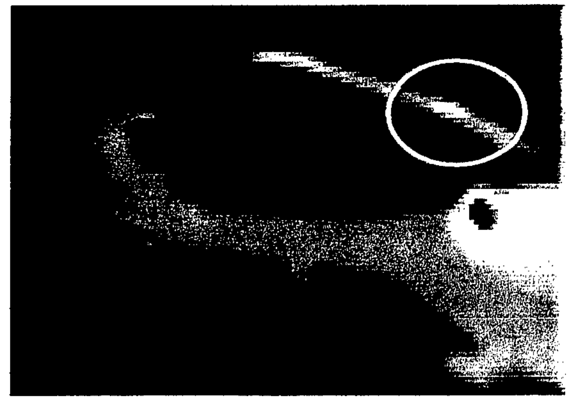


Figure 2Measurement IR image showing a strong plume reflection on the rotor structure

Dstl has recently investigated the ability to model such reflections through the use of a separate IR prediction system known as McCavity. McCavity is an IR solver for a discrete target. The system has been specifically designed to import unstructured CFD solutions and hence can take full account of platform plume and thermal characteristics. McCavity, developed for Dstl by Insys Ltd, combines advanced ray-tracing algorithms with a statistical Monte Carlo approach to simulate complex multiple reflections. This technique enables the interactions between thermal emissions, radiative heat transfer and surface-to-surface reflection phenomena to be predicted accurately. A flow diagram of the modelling process is summarised in Figure 3, indicating the role of McCavity with respect to the other codes.

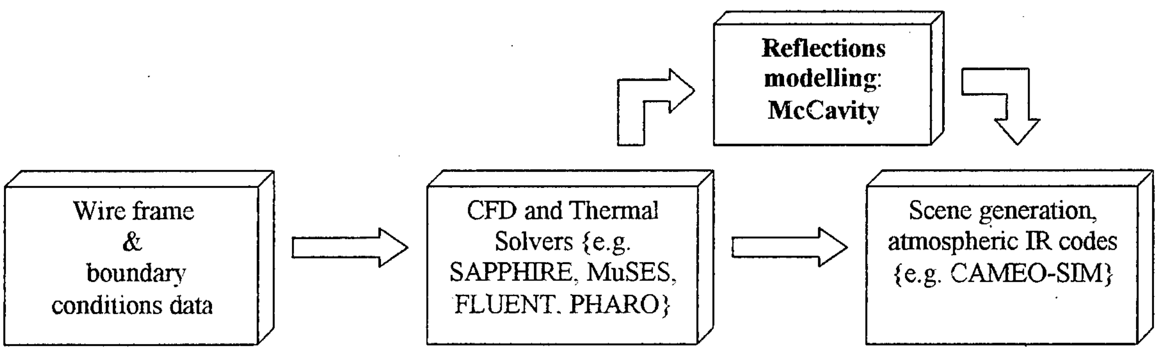


Figure 3 Flow diagram showing McCavity with respect to other modelling codes

This paper provides a brief introduction to McCavity, and the ongoing validation process is discussed. Validation against two standard published results is presented. These along with other similar simple validation test cases are providing a high degree of confidence for application of these techniques for IR signature reflection assessments. As a capability demonstration, the reflections arising from the exhaust plume of a generic helicopter are modelled and presented.

THE McCAVITY CODE

McCavity uses a Monte Carlo stochastic radiation prediction algorithm to determine the properties of the infrared signature for a discrete target. The Monte Carlo method is not discussed in detail as it has been extensively covered elsewhere [6, 7, 8]. The target to be modelled is imported with appropriate thermal and optical properties assigned to each of the individual facets either as a surface description (wire-frame) or a volumetric grid (as produced by CFD codes). The optical surface characteristics are defined with reference to the bi-directional spectral reflectance distribution function (BRDF) and/or the spectral hemispherical directional reflectance (HDR).

In general reflection modelling is dependent on the accurate calculation of view factors, F_{IJ}, between facets and gas cells. The view factor is generally defined as [8] the fraction of radiant energy leaving surface I of area, A_I, that is intercepted by surface J of area A_J, where,

$$A_I F_{IJ} = \int_{AI}\int_{AJ} \frac{\cos\theta_I \cos\theta_J}{\pi r_{IJ}^2} dA_I \, dA_J \quad (1)$$

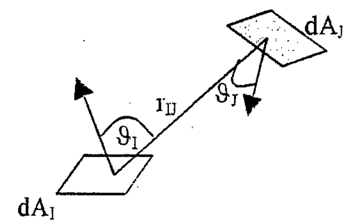


Figure 4 Schematic of view factor between two surfaces

However, this exchange term requires a directional factor to be introduced to account for non-lambertian reflections and emissions, as described in general terms by the BRDF function (which in a non-spectral form would be a 4-D function). In addition, facet to facet shadowing will need to be taken into account in some geometries. The consequent mathematical expressions are, except for simple cases, intractable for any analytical solution. Accordingly the favoured approach is to adopt a numerical integral.

In some cases, codes have been developed to solve using a radiocity solution with a fixed number of bounces. Here a Monte Carlo solution has been adopted which ensures that the appropriate number of reflections is fully accounted. This is particularly important for enclosed systems such as cavities, where many bounces are required to form an accurate solution. There are several different approaches to such problems, including analytical techniques, projection methods as well as Monte Carlo ray tracing algorithms. Emery *et al.* [9] provides a discussion of these in terms of potential accuracy and calculation times and concludes the Monte Carlo method probably represents the best numerical method for computing view factors for complex surface interactions. A study for Dstl conducted by Rolls Royce [10] drew similar conclusions, which led to the specification for McCavity.

The prediction method involves the generation of a statistically large number of random rays simulating the emission of photons either to or from the facet and gas cells. To account for directional emission for the solid surfaces the simulated emission has a weighted directional function applied to the random direction generator of the rays. Similarly the reflection properties are weighted such that the overall integral of a large sample of random rays will converge to the BRDF function. The gas cell radiation is treated isotropically. These rays are individually tracked using conventional geometric ray tracing techniques. The evolution of each ray is determined through another statistical assessment at each interaction for either an emitted or reflected energy contribution. The individual ray trace process halts when either the ray is absorbed by a facet or strikes an element of the image plane.

CODE VALIDATION

Validation of McCavity [11] has been an integral component of code development through rigorous testing at every stage. The code has been compared with several analytical solutions in the open literature to provide fundamental confidence in the computational accuracy. The accuracy of calculations performed by McCavity are controlled by three factors:

(i) The number of rays employed,
(ii) The numerical precision of the calculations, and
(iii) The accuracy of the input conditions of the problem under investigation.

The accuracy to a certain extent can be refined through increasing the number of rays employed and the detail within the initial model; however, this can drastically increase the processing requirements. There is always a compromise between confidence in the results and sensible processing times. Nominally accuracy at the 98-99 % level is expected from the Monte Carlo method with a practical number of rays.

Two examples presented here using simple cavity geometries with diffuse optical properties. The accuracy in the geometric representation is high, and therefore. the numerical precision of McCavity is under investigation. The published analytic solutions were programmed into a test harness for comparison purposes with McCavity. After obtaining a high degree of confidence in the test harness McCavity was run to the same input conditions with favourable results.

Sparow and Jonsson [12] examined the radiant emission characteristics of diffuse conical cavities and presented the efflux of radiation from a cavity opening. The apparent hemispherical emissivity (AHE) for a conical cone with angles ranging from 25 to 180 degrees is presented for emissivities ranging from 0.3 to 0.9. Figure 5 shows the geometry and a comparison between the McCavity results (symbols) and the analytic based test harness (line). They agree to within a 1.5 % error margin and for angles greater than 40° this is within 0.5 %. An error on the order of 2% is expected due to the statistical nature of the McCavity analysis method.

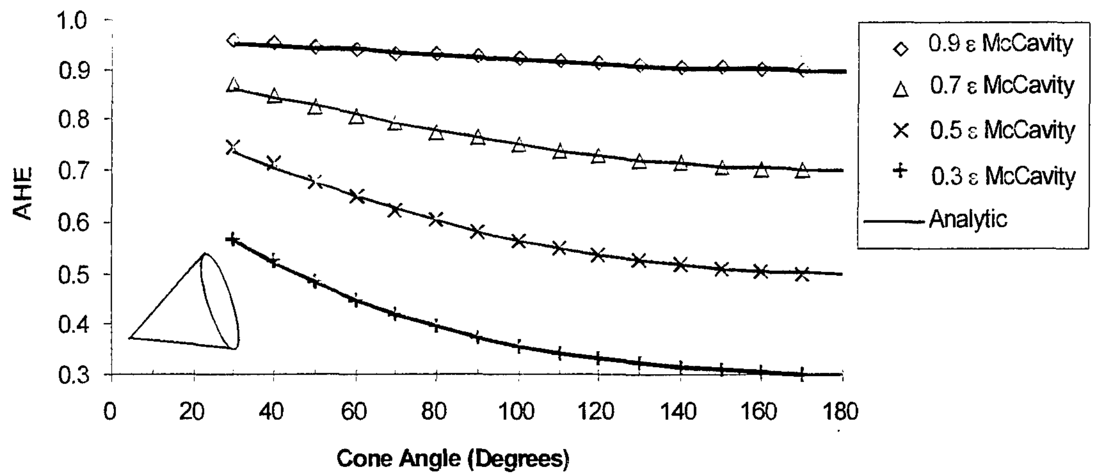


Figure 5 Comparison of apparent hemispherical emissivity (AHE) for an analytical solution [12] of a diffuse conical cavity with McCavity at emissivities of 0.3, 0.5, 0.7 and 0.9

Additionally, the apparent emissivity distribution along the cone surface was compared for a cone with semi-angle of 45 degrees and emissivity of 0.5; the results are shown in Figure 6. There is very good agreement between the analytical solution and McCavity, with the exception of points near the apex. This is considered to be due to an instability in the mathematics of the analytical solution, McCavity is not susceptible to this.

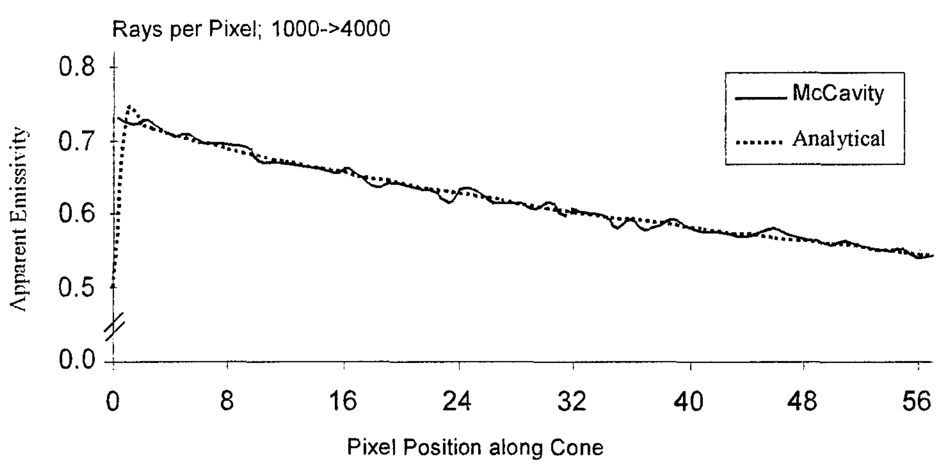


Figure 6 Comparison of Apparent Emissivity obtained analytically [12] for a diffuse conical cavity with McCavity at emissivtiy 0.5 and semi- angle 45 degrees

Similar comparisons have been made with Bedford and Ma's [13] analytic calculations, who examined the emissivities of diffuse cavities, for isothermal and non-isothermal cylindro-cones. A geometrical representation of a cylindro-cone, which consists of a cylinder with parallel sides and a conical end section, is shown in Figure 7.

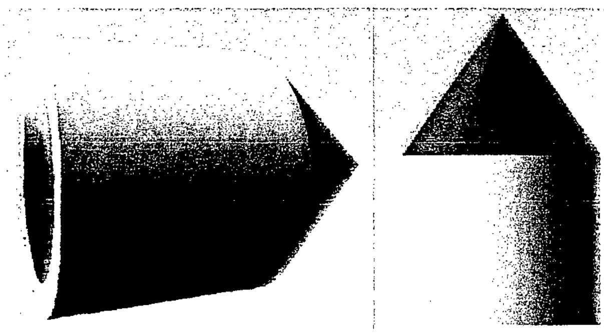


Figure 7 Geometric representation of a cylindro-cone [13]

The half-angle of the conical end piece was varied and the apparent hemispherical emissivity obtained both analytically as detailed in [13] and with McCavity, the results are shown in Figure 8. The agreement was seen to be excellent with less than a 1 % difference between the analytic solution and McCavity.

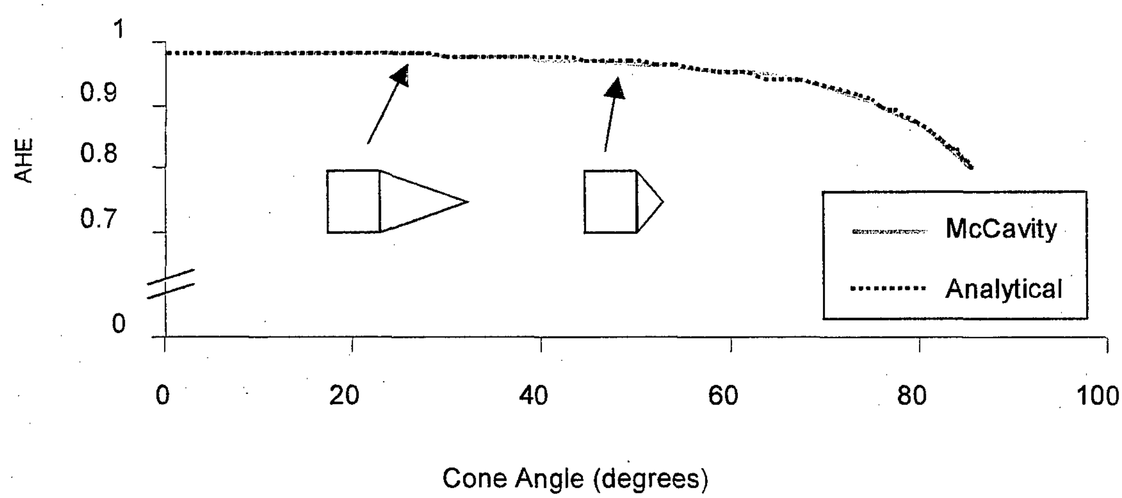


Figure 8 Comparison of apparent hemispherical emissivity (AHE) obtained analytically [13] with McCavity for increasing cone angle at an emissivity of 0.5.

Repeating the McCavity calculations would yield a normal distribution of values. In the presented examples the mean value converges toward the analytic solution. Additionally increasing the number of rays is expected to produce similar results in closer agreement to the analytic solution. The agreement shown here between the analytic solution and McCavity provides confidence in the McCavity algorithm. For more complex geometries and volumetric grids the accuracy is restrained by the input fidelity and the number of rays employed for a functional computation time. The optimisation of these are part of an ongoing validation programme.

EXTENTED SOURCE REFLECTIONS

As a capability demonstration, McCavity has been used to highlight the influence of reflections arising from a helicopter exhaust on a tail-boom and rotor structure. The three-dimensional CFD geometry generated for this study is shown in Figure 9, where the mass inlet boundaries used to represent the exhaust and down-wash are also indicated. Meshing was performed using a tetrahedral mesh with a sizing function based around the exhaust inlet. For this assessment arbitrary flow conditions were modelled. A fixed down wash velocity at (16 m s^{-1}), corresponding to a mass flux of 19.6 kg m^{-2} s^{-1}, was used to produce impingement of the exhaust plume onto the boom structure. The exhaust was defined with the initial boundary conditions representative of a 'top hat' exhaust gas at velocity 95.4 m s^{-1} (i.e., mass flux, 117 kg m^{-2} s^{-1}) and temperature, 596 K. Appropriate values for the principle IR emitting gas species, CO_2, H_2O, within the plume were also applied to the exhaust inlet.

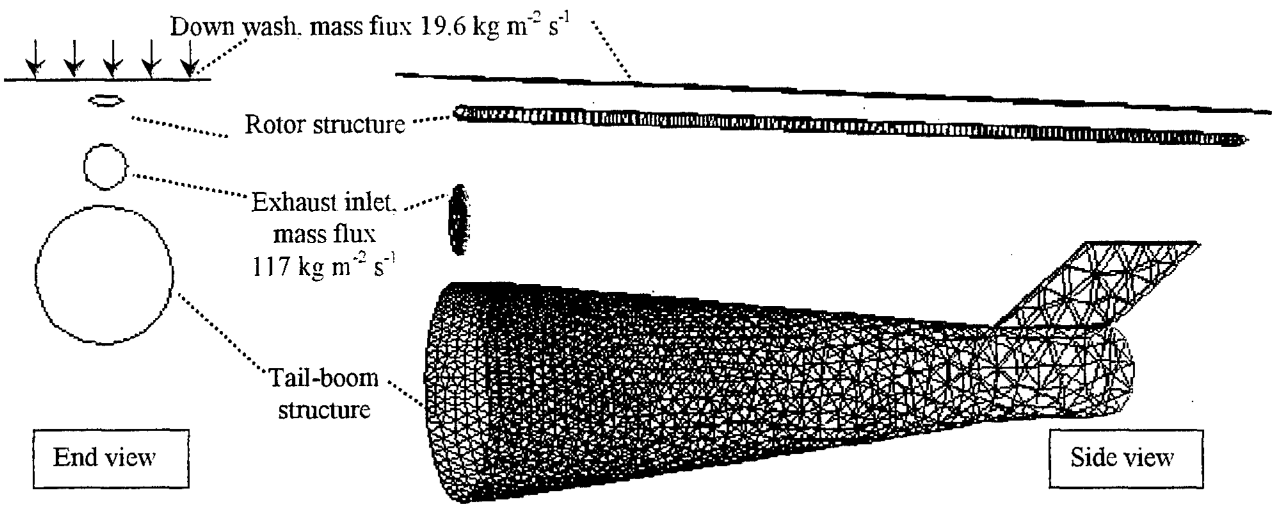


Figure 9 End and side view of a representative helicopter tail- boom and rotor structure with mass inlet boundaries shown

The CFD analysis was performed using FLUENT 6 with the K-ε turbulence model. The rotor flow was treated as a simple down-wash inlet with a static rotor blade defined just below this. The results obtained with the solver are presented in Figure 10. Image (a) shows the gas temperatures of a slice taken through the mid-section of the y-axis, demonstrating the plume clearly being directed down over the tail-boom structure. Image (b) shows the velocity vectors for this same slice; a concentration occurs in the proximity of the rotor structure due to a more detailed mesh. The temperature profile on the solid surfaces is shown in images (c) and (d) for two different views. Heating is experienced on the tail-boom with the majority occurring towards the rear of the structure. The final image (d) shows the underside of the rotor blade structure. There is no heating in evidence on the blade, as expected, since this component experiences no impingement from the plume.

Results from the CFD analysis were subsequently imported into McCavity to perform a reflection assessment. The material optical properties were defined as a 50 % diffuse, grey reflector uniformly across all the solid surfaces. The calculation was performed for the wavelengths 4 to 5 microns. Plots of radiant intensity [W sr^{-1}] output by McCavity are shown in Figure 11. The first image (a) shows the overall signature including contributions from both the plume and solid surfaces. The plume is visible due to the radiated emission of the hot species within it. The impingement-heated component of the tail-boom is also clearly visible. In addition to these two sources a reflection component is also included in this total image. These components are also presented separately for clarity. Image (b) shows the signature without the plume contribution. The plume component is presented in (c) and finally the reflected component only is shown in (d). Comparison between this reflected component and Figure 10(c) shows the addition of a high intensity region located along the top of the tail-boom, below the position of the exhaust plume. This demonstrates the contribution of reflections from the plume. A number of areas have been marked on these images and used for spectral analysis, which follows.

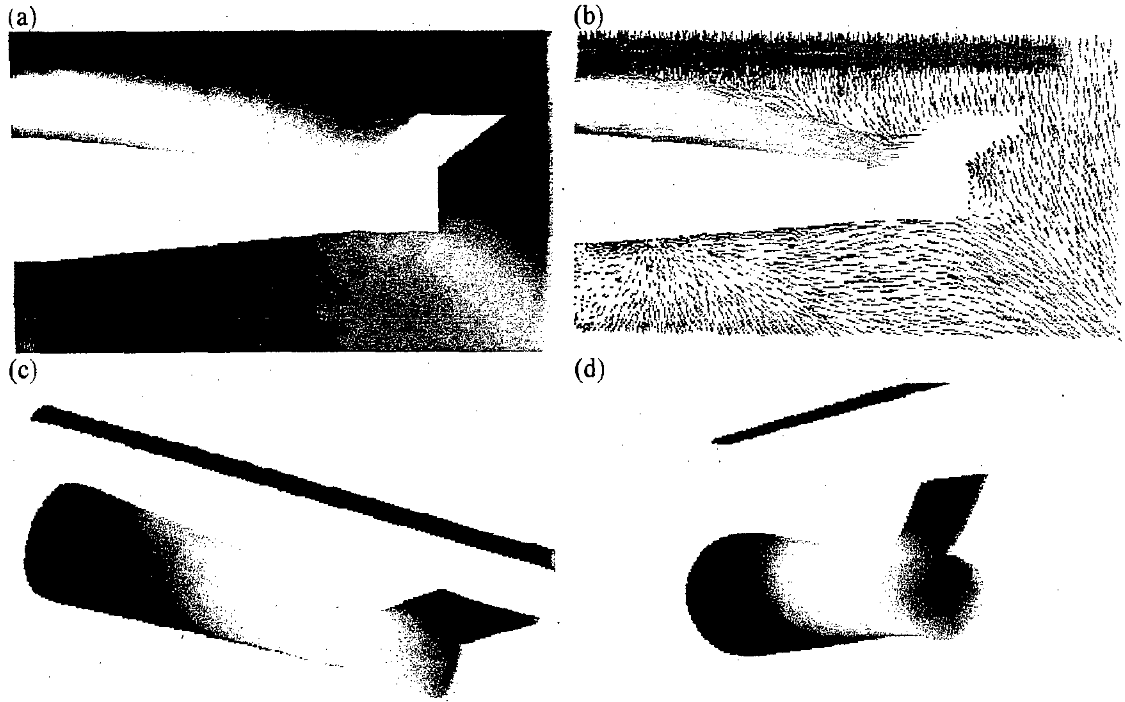


Figure 10 Temperature (K) profile of FLUENT CFD solution, (a) slice through central y-axis showing gas temperature, (b) central y-axis slice showing velocity vectors, (c) view down onto the helicopter tail-boom, (d) view up towards the underside of the rotor blade.

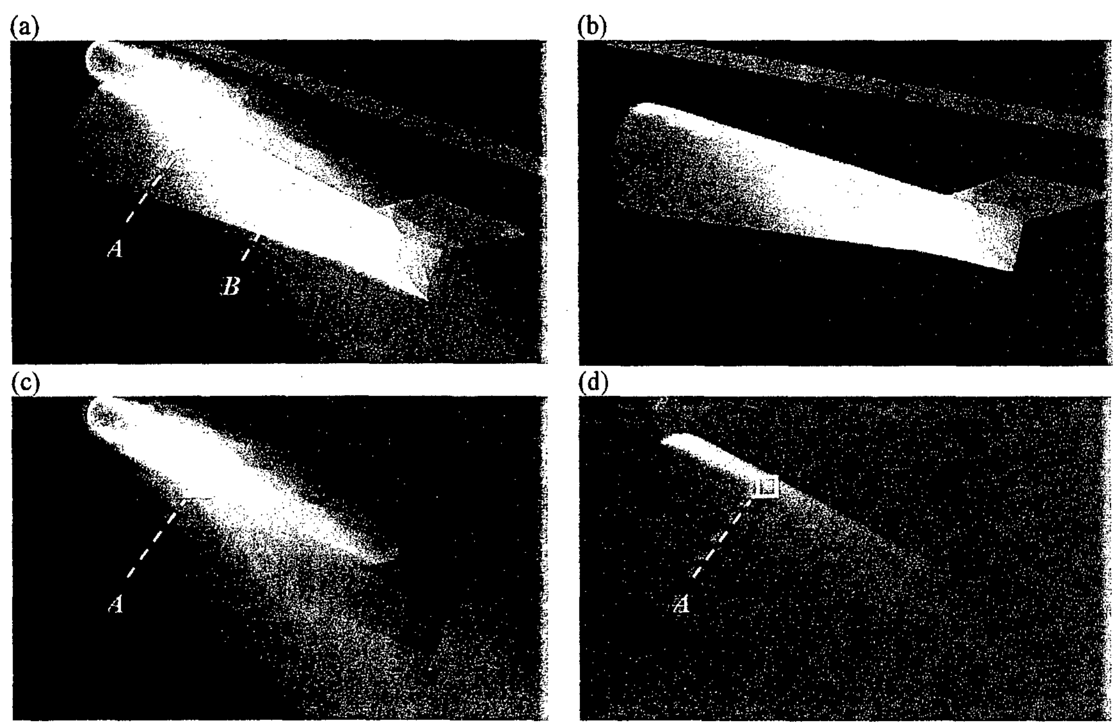


Figure 11 Radiant intensity images produced with McCavity ($W\ sr^{-1}$): (a) complete result, (b) view down onto the tail-boom with plume removed, (c) plume component only (d) reflection component only.

The spectral response obtained from the emission by a solid hot surface is expected to differ significantly from both the emission of a hot gas such as a plume and any reflected component of this. The reflected component would image the emission from the hot plume species at a reduced intensity determined by the surface reflectivity property.

To investigate the characteristics of the new high intensity region on the boom, the spectral response was taken between the wavelengths 4 and 5 microns for the regions shown in Figure 11, areas *A* and *B*. These were chosen to have the same image area of 91 pixels translating to an estimated target area of 2.08 m^2. *B* is located within the impingement-heated boom structure. *A* is taken from a region where both there is both a plume and reflection component, considered independently. The result is presented in Figure 12.

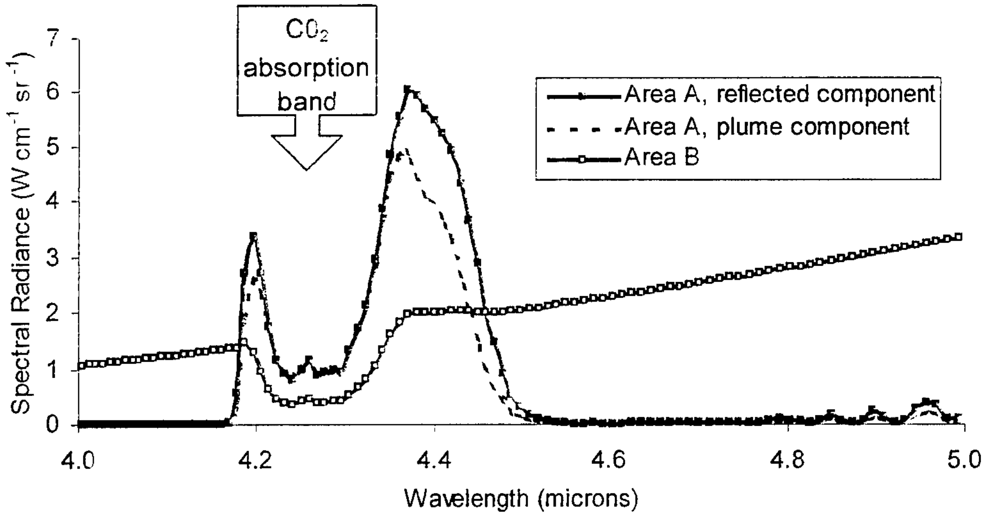


Figure 12 Spectral radiance of selected regions between the wavelengths of 4 and 5 microns.

A typical grey-body emission is obtained from area *B* as expected. This response has two contributing components, the presence of cool black body radiation, reduced by the emissivity of the surface, and some absorbance due to the species CO_2. This manifests as a dip in intensity between the wavelengths 4.20 μ to 4.37 μ. Both the reflected and plume component within area *A* show a completely different response. Significant radiance occurs only between the specific wavelengths of 4.17 μ and 4.50 μ, indicating the absence of continuum emission. This spectrally selective emission is characteristic of CO_2 fundamental (anti-symmetric stretch) gas emission. The signature arising from solely the reflected component is larger than the radiance from the emitted component. This has been calculated from the image with the plume removed. Consequently, there is no absorption of this signature. However, it demonstrates the potential for a significant reflection contribution even for a 50 % reflective, diffuse material.

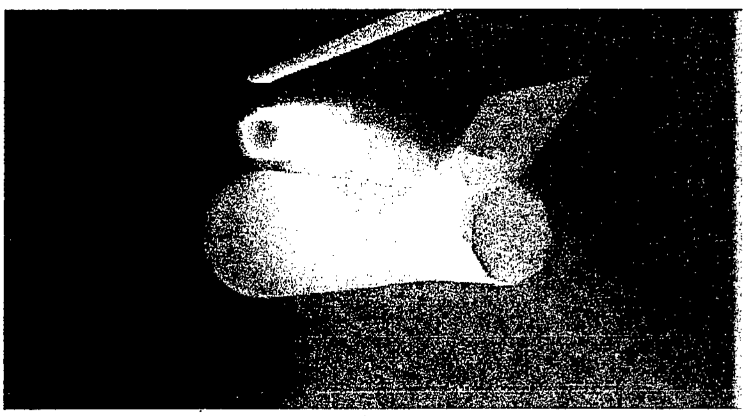


Figure 13 McCavity radiant intensity image (W sr^{-1}) showing a view of the rotor blade underside.

The image in Figure 13 takes a look-up view towards the underside of the rotor blade. This shows evidence of a signature component along the lower surface of the rotor blade. Comparison of this with Figure 10(d), which shows this thermal contribution was not previously present, therefore, confirms this as a plume reflection. This reflection would not suffer absorption from the plume due to its displaced position, potentially providing a high source of radiance.

This simple modelling example has successfully demonstrated the capability of McCavity to simulate reflections arising from the heated exhaust from a helicopter powerplant. Realistic structures are likely to be far more complex with many components acting as reflecting surfaces. The reflection contribution is highly dependent upon the target view, particularly for non-diffuse surface properties. Therefore, to fully assess the implications is a non-trivial exercise.

CONCLUSIONS

A Monte Carlo based ray tracing code, McCavity, has been developed, which can be used for modelling the reflection contributions to the infra red signature of targets.

Confidence has been shown in the calculation procedure through the successful comparison with published analytically derived solutions. The agreement level was well within the 1-2 % tolerance nominally expected from Monte Carlo processes for useful run times.

The application of McCavity to model reflection contributions from an extended hot gas source has been successfully demonstrated with a generic helicopter tail-boom and rotor structure using an unstructured CFD analysis. The importance of modelling extended source reflections has been shown and in some instances, such as rotor blade reflections, these have the potential of being significant.

ACKNOWLEDGMENTS

This work was sponsored by assignment 1221 under the UK MOD Defence Research Package for Tri-Service Helicopter Research.

REFERENCES

[1] S Rooks, A.G. Smith, "Flow field and IR Signature Modelling for Ground Targets", *Proceedings of the Eighth Annual Ground Target Modeling & Validation Conference,* Houghton, MI, August 1997

[2] N Lavery, S Brown, J Spittle, S Rooks, L. Hayward, A.G. Smith, "Improved modelling capabilities of PHARO – a coupled CRD-Thermal solver for detailed IR prediction", *Proceedings of the Twelfth Annual Ground Target Modeling & Validation Conference,* Houghton, MI, August 2000

[3] D Oxford et al, "CAMEO-SIM: A physically accurate broadband EO scene generation system for the assessment of air vehicle camouflage schemes", *Proceedings of the Ninth Annual Ground Target Modeling & Validation Conference,* Houghton, MI, August 1998

[4] M.A. Gilmore et al, "Synthetic scene simulation for camouflage assessment", *Proceedings of the Tenth Annual Ground Target Modeling & Validation Conference,* Houghton, MI, August 1998

[5] D. Oxford, D. Filbee, C. Stroud, "CAMEO-SIM: Development history and example applications", *Proceedings of the Eleventh Annual Ground Target Modeling & Validation Conference,* Houghton, MI, August 2000

[6] R. Siegel, and J.R. Howell. "Thermal radiation heat transfer", 3ed, Hemisphere publishing corporation, 1992

[7] J.R. Howell. and M. Perlumutter, "Monte Carlo solution of thermal transfer through radiant media between gray walls", *J. Heat Transfer*, pp 116-122, Feb 1966

[8] R.C. Corlett. "Direct Monte Carlo calculation of radiative heat transfer in a vacuum", *J. Heat Transfer*, pp 376-382, Nov 1966

[9] A.F. Emery, O. Johansson, M. Lobo. A. Abrous. "A comparative study of methods for computing the diffuse radiation viewfactors for complex structures". *J. Heat Transfer*, Vol 113, pp 413-422, 1991

[10] A.T. Jones, Classified Report, Rolls-Royce Applied Science Laboratory, [RR(OH)1461], 1996

[11] D. Brunnen, "Verification of McCavity against open literature sources", Hunting Engineering, Reddings Wood, Bedford, MK45 2HD, UK, [HE/336/TR/008], 2001

[12] E.M. Sparrow and V.K. Jonsson, "Radiant emission characteristics of diffuse conical cavities", *J. Optical Soc. Of America*, Vol 53, No 7, pp 816-821, 1963

[13] R.E. Bedford and C.K. Ma, "Emissivities of diffuse cavities, II: Isothermal and nonisothermal cylindro-cones" *J. Optical Soc. Of America*, Vol 65, No 5, pp 565-572, 1975

Army Aviation FLIR Mission Planning Enhancement

George G. Koenig, Robert E. Davis and Major Stephen T. Milton
U.S. Army Engineer Research and Development Center
Cold Regions Research and Engineering Center
72 Lyme Road, Hanover, NH 03755

ABSTRACT

Aviators rely heavily on Forward Looking InfraRed (FLIR) imagery to navigate and to rapidly and accurately detect and identify targets. Weather and weather impacted terrain and targets can significantly alter terrain-target infrared contrast relationships impeding an aircrew's ability to effectively use FLIR imagery for mission accomplishment, potentially increasing aviator's exposure to enemy threats and counterattacks, and ultimately decreasing system lethality and survivability. Pre-flight physics-based "through the sensor" infrared synthetic scenes can mitigate the impact of weather by portraying the weather impacted terrain-target infrared battlefield scenes accurately.

The U.S. Army Engineer Research and Development Center (ERDC), the Air Force Research Agency (AFRA), and the U.S. Army Aviation Test Directorate tested the utility of the Infrared Target-scene Simulation Software (IRTSS) system as a mission planning and rehearsal tool for Apache attack aviation. Pre-flight infrared synthetic scene mission enhancement was quantified based on Battle Position (BP) rankings, as compared to the rankings of a Standardization Instructor Pilot (SIP), target detection times, number of target false detects, and number of target non-detects. Questionnaires were used to qualitatively assess the 'value' of pre-flight synthetic infrared scenes as judged by Army aviators.

Predicted FLIR scenes significantly enhance the Intelligence Preparation of the Battlefield (IPB) process, allowing aviation planners to predict FLIR performance, evaluate and select optimal routes, and battlefield positions. Two groups of fifteen Apache pilots participated in the test to determine the impact of infrared synthetic scene simulation on Army aviation mission effectiveness. The control group received the current pre-flight briefing tools, while the test group received current briefing tools plus IRTSS scenes representing the Apache Target Acquisition Detection Site (TADS) imagery consistent with the mission Fragmentary Order. IRTSS scenes improved pilot performance in all test areas. Battle Position selection improved by 75%, target acquisition by 61%, target detection by 41%, and time to detect decreased on the average 6.5%.

INTRODUCTION

Army aviation warfighters rely heavily on Forward Looking InfraRed (FLIR) sensor systems to quickly and accurately acquire and identify targets. FLIR imagery is used for visual cueing for navigation, identifying Battle Position/Firing Position (BP/FP), and orientation in the Engagement Area (EA). FLIR imagery provides a "visual perception" of the battlespace thus enhancing two key enablers of the Objective Force: information dominance and situational awareness. Rules of Engagement (ROE) require Army attack aviators to visually recognize their target before releasing their ordinance. Even under blue-sky conditions, complex spatial and temporal variations in the thermal signatures of natural backgrounds and targets complicate the "visual recognition" of targets using FLIR systems. But, FLIR systems provide a clear and unequivocal advantage for nighttime engagements, and can even provide an advantage over 'looking out the window' during the daytime. Target detection is a contrast radiant intensity issue. In general, the greater the contrast between the background and the target the easier it is to detect the target. In the visible, paint and paint patterns have been developed to basically reduce the contrast between the target and the background. It is more difficult to mask 'hot spots' on vehicle targets thus FLIR systems can offer a clear advantage over optical systems. Weather impacted terrain and terrain-target contrast relationships affect FLIR mission performance (Bryant, 1998). Conditions that degrade FLIR system performance can increase aircraft exposure to enemy threats and counterattacks, increase the time required to detect and identify targets, and increase the time to accurately select BPs/FPs and orienting on EAs, ultimately decreasing system lethality and increasing system vulnerability. Generally, the less contrast between background terrestrial features and between the target and the background, the longer it takes a human to make confirmed recognition (Bryant, 1998). Today, the Army does not have a tactical level capability to predict

FLIR scenes of the battlespace. Warfighters must rely on their ability to mentally translate two-dimensional topographic maps and visual animations to the corresponding FLIR representation of the battlespace. Unfortunately, a one-to-one mapping between visual and corresponding infrared imagery does not exist. Achieving situational awareness in FLIR is a most challenging endeavor (Milton and Williams, 2002). Synthetic infrared scenes generated using physic-based models have the potential, when used as a pre-mission briefing tool, to significantly enhance FLIR situational awareness by allowing aviators to view navigation routes, BP/FP positions, EA appearance, and target-background thermal contrast based on expected environmental conditions at mission time. A Concept Experimentation Program (CEP) was conducted to assess the military utility and value pre-mission, synthetic IR scenes provide to the aviation warfighter.

CONCEPT EXPERIMENTATION PROGRAM

The Concept Experimentation Program provides the Army Training and Doctrine Command Battle Laboratories a method to evaluate and capitalize on emerging technologies, material initiatives, and warfighting concepts while offering the research and development community an effective and efficient method of determining the value added to warfighting capabilities. The Air Maneuver Battle Laboratory and the Aviation Directorate of Combat Developments at Fort Rucker, AL, sponsored a CEP collaborative effort involving researchers at the Engineer Research Development Center (ERDC/Cold Regions Research Engineering Laboratory and ERDC/Topographic Engineering Laboratory) and defense contractors supporting the Air Force Research Laboratory (AFRL) to determine the military use and benefit of predicted IR scenes of the battlespace on Army attack pilot performance. Conceptually, the experiment was done in the context of the tactical operating domain of AH-64A Apache helicopters employing hellfire missiles for high priority targets. The AH-64A uses a Target Acquisition Designation Site thermal sensor, operating in the 8-12 micrometer band, for target detection. This CEP focused on the military worth of predictive TADS scenes and animations and specifically addressed the following issues:

Issue 1. Battle Position Evaluation: Do pre-mission synthetic IR scenes improve the AH-64A pilots' ability to evaluate and rank order BPs?
Issue 2. Target Detection and Identification: Do pre-mission synthetic IR scenes decrease the time it takes to detect and identify targets? Do pre-mission synthetic IR scenes improve target detection (decrease the number of false detects) and improve target acquisition (decrease the number of non-detects)?
Issue 3. Situational Awareness and Risk Mitigation: Does the capability to generate IR scenes from a "look-back" position increase situational awareness thus enhancing risk mitigation?
Issue 4. Enhance Aviation Mission Planning: In the opinion of the test subject would access to pre-mission, synthetic IR scenes (e.g., through Aviation Mission Planning System-AMPS) improve the Intelligence Preparation of the Battlefield process, Battalion-Platoon planning/unit rehearsals and aircrew/aircraft risk mitigation?

The CEP experiment was designed to obtain quantitative and qualitative measurements using IRTSS generated "through the sensor" synthetic images, in a classroom setting, to address and answer the issues stated above.

THE EXPERIMENT

The classroom experiment was conducted July 9-13 2001, at Fort Hood, TX, and was administered by the Aviation Test Directorate (AVTD), U.S. Army Operational Test Command to determine the value added to mission planning and rehearsal of synthetic TADS scenes and animations. The experiment consisted of 30 Apache pilots crossed-leveled by flight experience and placed into two groups: the baseline group and the IRTSS group. Both groups contained company-level officers ranging from WO1 to captain. Both groups received the standard mission planning and rehearsal tools (operations order, topographic map, operational overlay, and Aviation Mission Planning System and bold earth line-of-sight application). The IRTSS group received 'through the sensor' IRTSS predicted TADS scenes and animations. All other aspects of the experiment for the two groups were the same, including the level of training and time to complete the test. The IRTSS group was told the "The IRTSS scenes and animations closely represented what they would see if they actually flew the specific missions as outlined in the operations order". In order to address issues 1 Apache TADS video was collected from 5 BPs associated with two separate EAs. Each EA contained three VISMOD HMMWVs. The same approach was used to address issue 3, but TADS video was collected for four different EAs and a total of 8 BPs. Each EA contained a single VISMOD HMMWV. The Apache TADS IR video was obtained for: 1) the helicopter unmasked to view an area potentially containing target(s), and 2) scans of an (notional) EA containing targets. Pilots were tested individually under the supervision of the AVTD test administrator. A personnel computer was used to view all TADS video and IRTSS synthetic scenes and animations. The pilots from the two groups (baseline and IRTSS) were given the same FRAGmentary Order (FRAGO).

Upon completion of the review of the FRAGO, the pilots had access to their groups' respective mission planning tools. The AVTD test administrators conducted the tests by instructing the pilots to watch a series of pre-recorded and digitized TADS FLIR video. The test did not evaluate the IRTSS GUI or the validity of the IRTSS generated TADS synthetic scenes or animations.

SYNTHETIC SCENE GENERATION MODEL: IRTSS

Physics-based modeling systems used to generate synthetic IR scenes must have the capability to predict the radiance at the aperture of an IR sensor, including the impact of atmospheric transmission, over the spectral response band of the sensor. This entails thermal models that predict the energy and mass transfer processes, atmosphere-surface interactions, and suitable computer architecture. Figure 1 provides a notional flow of information from the geophysical data bases required for model initialization to rendered synthetic scenes. We used a modified version of the Infrared Target-scene Simulation Software (IRTSS) developed by Radex Corporation under Air Force Research Laboratory (AFRL), Hanscom AFB, MA (Seeley and Luker, 1998) sponsorship. Because of limitations imposed by the sensitive nature of performance specifications of military sensor systems, the TADS sensor system was modeled using a top hat response function over the spectral interval from 8 to 12 micrometers.

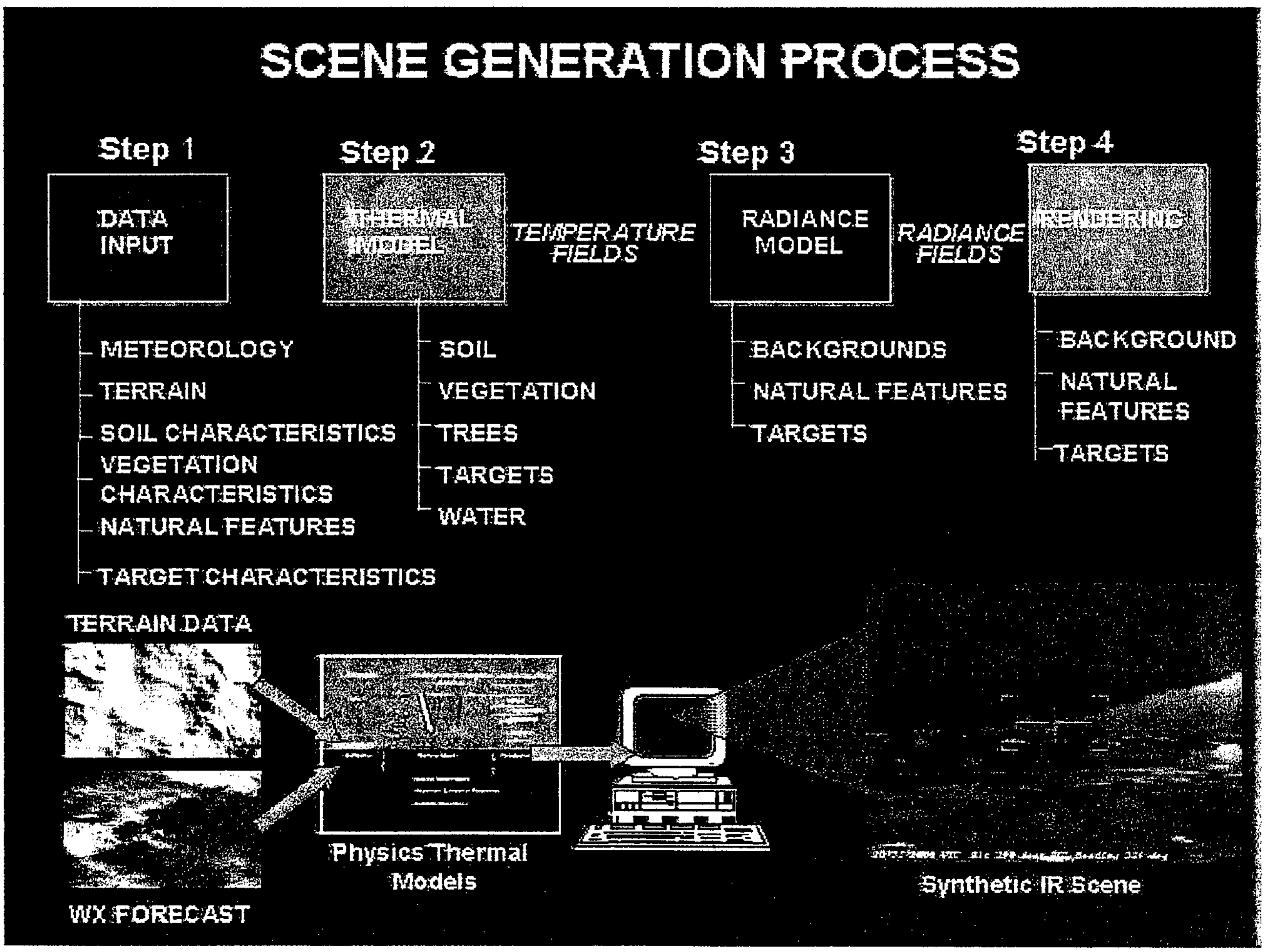


Figure 1. Notational chart of the process to predict IR scenes. Four distinct modules involved in the process: model input data, predicting the temperature fields, generating the radiance fields, and rendering the synthetic IR scene.

Geophysical data requirements for synthetic scene generation

The process of generating synthetic IR scenes starts with the specification of the geophysical data associated with the location to be simulated. The geophysical information required consists of: digital terrain data, terrain feature data, and the physical, thermal and optical properties associated with the feature data. The digital terrain data for the Fort Hood experiment was obtained from various Geographic Information Systems, including those at CRREL, TEC, the Fort Hood Integration Training Management (ITAM) office, and Pacific Meridian Corporation for an area of 20 by 20 km with a spatial resolution ranging from 1 to 3 meters. The terrain elevation data is used to calculate the slope and aspect of the terrain, at each elevation posting, required to correctly ascertain solar loading. The feature information provides the surface soil type, sub-soil type, surface vegetation, and the thermal, optical and physical properties associated with the terrain feature data. The vegetation can be: no vegetation-bare soil, grass, or canopy. Canopy, information consists of: canopy type (deciduous or coniferous), canopy density, and basic canopy physical information (tree height and crown diameter). The canopy information had insufficient spatial resolution to support the CEP synthetic scene generation requirements. In general, the Apache pilots indicated it was not necessary to model every tree and branch, but it was necessary to have the location of tree clusters notionally correct. The canopy feature data missed many of the individual tree clusters and in areas of more continuous canopy cover missed gaps in the canopy that were large enough and interconnected to the point where it was possible to move military vehicles through the canopy. To rectify this problem, CRREL developed semi-automated, interactive techniques to add and remove trees from the original canopy cover information using satellite imagery.

CRREL installed an automated weather station near the Jack Mountain training and maneuver area. The weather station was operated continuously from 1 October 2000 until 30 April 2001. The weather data collected consisted of total and diffuse downwelling solar radiation, upwelling solar radiation used to compute albedo, upwelling and downwelling infrared radiation, air temperature, relative humidity, wind speed and direction, precipitation amount and rate, and IR thermometer measurements of surface temperature. The meteorological information is used to initialize IRTSS and as boundary conditions for generating the synthetic TADS scenes. A minimum of five days of weather information is required to "spin up" IRTSS to achieve thermal stability before synthetic scenes can be generated.

Thermal and Thermal Radiation Models

The thermal models employed in IRTSS were developed under the Joint Test and Evaluation Smart Weapons Operability Enhancement (JT&E SWOE) program (Welsh, 1994: Koenig et. Al. 1995; Welsh and Link, 1995). The SWOE models, and IR synthetic scenes produced using the SWOE models have undergone significant validation (Welsh, 1994; Siegel and Castellan, 1988). The results of the validation of the SWOE models indicate absolute accuracies on the order of a few degrees Kelvin and relative accuracies on the order of one-degree Kelvin. IRTSS uses SWOE models for bare and snow covered ground, forest areas, and vegetation over ground. The physics-based thermal models predict the surface temperature for a series of relatively homogeneous polygons with uniform thermal properties and surface features (Kress, 1992; Ballard, 1994). Homogeneity is based on the slope and aspect of the terrain, soil type, and surface vegetation, i.e. grass vs. forested areas. The soil/snow thermal model is based on the work of Jordan (1991), which simulates most of the important physical processes in snow, but assumes that conduction dominates heat transfer in the soil. The vegetation model follows the approach of Balick et al (1981) and is coupled with the soil/snow model. Solar and infrared radiation interactions between the soil surface and the overlying vegetation are modeled, but physical processes like root zone moisture uptake are not. The canopy model follows the scheme originated by Verhoef and Bunnik (1975) and extended by Smith (1981). Smith used expressions for the canopy energy and mass balance assuming a plane parallel five-layer canopy model (three canopy layers, an atmospheric layer above the canopy, and an underlying soil layer) with uniform canopy properties. Again, the solar and infrared radiation interactions between the soil and the canopy are modeled. The original SWOE model suite included a series of models to compute the spectral thermal radiation associated with natural terrain features in the spectral region from approximately 1 to 20 micrometers. These models included emitted thermal radiation and primary and secondary reflection of skyshine and thermal emitted radiation from scene elements. Because of the computation expense, IRTSS adopted a much simpler approach. Only the emitted thermal radiation is modeled using Stefan-Boltzmann's law. Target temperatures and the associated emitted thermal radiation are calculated using Multi-Service Electro-optic Signature (MuSES) (Rynes, et. al., 2000). The atmospheric transmission of the emitted thermal radiation from natural backgrounds and targets is handled using the DoD model standard for atmospheric transmission, MODTRAN (Berk, et. al., 1989).

Infrared Target Scene Simulation Software

IRTSS provides the capability to generate "through the sensor" predicted IR, night vision, and visible scenes and animations of weather impacted natural backgrounds and targets. Originally developed to support Air Force (high and fast) weapons

systems, IRTSS was modified to support Army Aviation (low and slow) Forward Looking Infrared TADS systems on Apache AH-64 helicopters. The Army version of IRTSS has added trees and structural features (buildings, etc.) to portray realistic line-of-sight obscurations that could potentially be encountered during an Army aviation mission. IRTSS provides the Army aviation warfighter with pre-flight awareness of the impact of weather and terrain as seen through the TADS system.

IRTSS runs under a client-server architecture on either a UNIX or LINUX based computer system. Scenario generation is achieved using a Graphical User Interface (GUI) that provides capabilities that increase situational awareness. For example, mask-unmask scenario generation, a look-back capability, terrain following fly through, and a 360-degree EA fly around to ascertain the best avenue of approach to an EA and best BP to unmask for target kill. For the CEP, IRTSS incorporated a sensor model representing a NATO FLIR system, with sensor fields of view and sensor characteristics similar to the Apache TADS. Therefore, the IRTSS generated scenes and animations represented a credible surrogate for the CEP test subjects, displaying what the Co-Pilot/Gunner would see on an actual TADS system flying the FRAGO. Figure 2 shows a comparison of an IRTSS generated scene to the corresponding TADS scene digitized from the actual video footage. Viewing geometry, location within the Fort Hood test range, time of day, and weather conditions are the same for the generated scene as the TADS video. The sequence of events to generate a synthetic scene is fairly straightforward. First, the user specifies a geographic location populated with targets from the databases packaged with the IRTSS system. Next, the user executes IRTSS using automated weather data, or locally collected weather information, as was the case with the CEP experiment. Using automated weather data from a mesoscale weather model (for example, MM5 or the IMETS/BFM) would provide a prognostic capability allowing the generation of synthetic scenes over the forecast period of the mesoscale weather model. Model execution takes several minutes. Once the thermal fields associated with the scenario have been generated, synthetic scene(s) can be generated using different ranges, altitudes, and headings without rerunning the thermal models. IRTSS presently comes bundled with several geographical data sets; over 50 different target geometries including buildings, bunkers, surface-to-air missile (SAM) sites, and wheeled and tracked vehicles; and several sensor models.

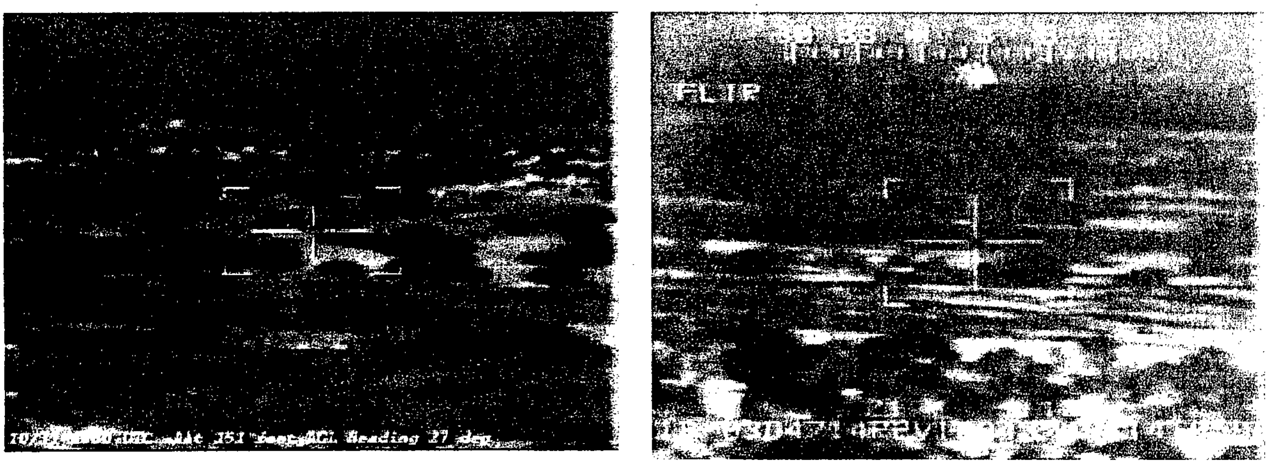


Figure 2. Comparison of an IRTSS scene (left panel) and actual AH-64 TADS FLIR scene (right panel) for the same location, time, and weather conditions. Both have medium field of view, white-hot polarity.

RESULTS

The Army Test and Evaluation Command (ATEC) conducted the formal analysis of the CEP data using a series of statistical and categorical procedures to directly address issues 1- 4. Key elements of the analysis conducted by ATEC have been extracted from the ATEC report titled "Test Report for Forward Looking Infrared (FLIR) Mission Planning Performance Enhancement", report 2001-AVN-1105, October 2001 and presented below.

Issue 1. Battle Position Evaluation: In order to establish a standard of reference, an aviation Standardization Instructor Pilot with access to all materials, including live pre-recorded TADS videos and IRTSS products, established the rank order of the five BPs for each EA. This ranking was considered 'truth'. As indicated earlier, the baseline group received the standard planning tools while the IRTSS group received both the standard planning tools, IRTSS scenes looking from the BP toward the EA, "look backs" from the EA to the BP, and animations of unmasking and scanning within the EA. The "look back"

provides the pilot with some idea of how he is silhouetted against the background. In general, the greater the clutter in the background the more difficult it will be to detect the aircraft. The standard planning tools consisted of: FRAGO, 1:50,000 topographic maps with the position of the BPs and corresponding EAs, Line-Of-Sight (LOS) plots from the BPs to the EAs on a 1:50,000 scale map from the Aviation Mission Planning System (AMPS) in increments of 50 feet from 50 feet to 200 feet elevation, and a copy of tasks, conditions, standards and other reference material. The test subjects had 20 minutes to review the material associated with each EA and were then asked to rank the BPs. ATEC used the Spearman Correlation Coefficient to determine the correlation between the BP rankings for the baseline group and the IRTSS group, and the SIP ranking of the BPs. At the 80% confidence level for EA Maim the BP rankings of 8 of the 15 pilots in the IRTSS group correlated with the SIP rankings while 4 of the 15 pilots in the baseline group correlated with the SIP rankings (see highlighted rows in table 1). This equates to a 100% improvement in BP rank order relative to the SIP (truth) BP rank order. At the 80% confidence level for EA Injure the BP rankings of 6 of the 15 pilots in the IRTSS group correlated with the SIP rankings while 4 of the 15 pilots in the baseline group correlated with the SIP rankings. This equates to a 50%improvement in BP rank order relative to the SIP (truth) BP rank order. Table 1 is the ATEC analysis. Post-test surveys showed 100% of the pilots in the IRTSS group indicated IRTSS scenes and animations helped in the selection of the optimal BP. Every IRTSS pilot indicated the IRTSS unmasking animation clearly showed the impact of vegetation (trees) on LOS and was an improvement over the AMPS bold earth LOS. For example, CPT G. Heap, Company Commander, 1-4 Avn. BN, 4th ID, stated "Accurate vegetation data provided more FLIR visual cues ..... makes pre-mission EA development easier...".

Table 1 Frequency Distribution of Spearman Correlation Coefficients by Battle Position

	BP MAIM		EA INJURE	
Coefficient Range	Baseline	IRTSS	Baseline	IRTSS
<0	3	1	1	1
0 – 0.2	4	3	1	2
0.21 – 0.4	1	0	3	4
0.41 – 0.6	3	3	6	2
0.61 – 0.8	2	5	3	1
0.81 – 1.0	2	3	1	5

Issue 2. Target Detection and Identification: After allowing 5 minutes for "mission planning", each test subject viewed sequentially 8 target detection scenarios. Each scenario consists of an average of 60 seconds of TADS video from a BP scanning an EA containing a single VISMOD HMMWV. For this test 4 EAs (EA shell-3 BPs, EA Bravo-3 BPs, EA Graham-1 BP, and EA Stampede-1 EA) and 8 BP were used. All test subjects previewed the same planning information and viewed the target scenarios in the same order. In addition to the IRTSS scenes, the IRTSS group viewed animations scanning the EA from the BP(s) associated with that EA. To minimize potential biasing of the IRTSS test subjects, the direction of scanning in the animations did not correspond to the direction of scanning on the TADS video. The metrics used to quantitatively address issue 2 consists of the time to detect, the number of false detections, and the number of failure to detect (non-detects). Test subjects were asked to indicate to the test administrator when they believed they detected the target in the TADS video. At this point the test administrator noted the time since the start of the video and asked the test subject to point to the location of the target on the video display. If the test subject correctly identified the target location the time to detect was recorded. If not, a false detect was noted and the video and timer started from the stop point. This procedure was repeated until the test subject correctly identified the target location or the video ended. If during the course of the test, the test subject did identify the correct target location, the number of false detects and the total accumulative time to detect was noted. If the subject did not identify the target location during the course of the test a non-detect or failure to detect was recorded. The procedure was followed for all eight scenarios. For each false detect the test subject was given a 5 second penalty. While this value is somewhat arbitrary, the normal mode of operation for AH-64 Attach pilots is to operate their TADS system in Medium Field Of View (MFOV) and to switch to Narrow Field Of View (NFOV) to confirm target detection. If target detection is not confirmed, they switch back to MFOV and continue to search for the target. This process can take on the order of 5 seconds and hence the 5-second penalty imposed for a false detection. The results of the ATEC analysis are presented in table 2. The number of target detection scenarios is 120 (15 pilots per group, times 8 BP). The BPs associated with the EAs have been numbered. Some of the BPs are used to view two different EAs. In this case, the BP number is the same for the different EAs. ATEV used the Student t-test at the 90% level of significance to test for differences between the baseline group and the IRTSS group. For the time to detect, only the analysis of EA Bravo indicated a significant difference between the two groups. While a penalty was assessed for false detects, no additional time penalty was added for non-detects. In the opinion of the pilots, the scenarios associated with EA Bravo were the most challenging.

Table 2. Battle Position and Engagement Area Mean Detection Times (Seconds)				
Engagement Area	Battle Position	IRTSS Mean Time	Baseline Mean Time	Statistically Significant Difference
Bravo	5	41.2	60.3	Yes
	6	33.5	38.2	No
	7	14.1	16.5	No
	All	29.6	38.3	Yes
Graham	3	20.5	19.3	No
	All	20.5	19.3	No
Shell	4	34.9	27.9	No
	5	18.7	18.5	No
	6	36.1	36.3	No
	All	29.9	27.6	No
Stampede	5	35.4	33.6	No
	All	35.4	33.6	No
Overall		29.3	31.3	No

The student t-test at the 90% significance level was also used to determine if there was a significant difference between the two groups on the number of false detection. The result of the ATEV analysis indicates there is a significant difference at the 90% significance level. Table 3 is a breakout of the false detects for the two groups by EA and BP.

Table 3. Number of False Detections by Battle Position and Engagement area			
Engagement Area	Battle Position	IRTSS	Baseline
Bravo	5	5	16
	6	1	7
	7	3	0
	All	9	23
Graham	3	0	0
	All	0	0
Shell	4	2	6
	5	0	0
	6	4	2
	All	6	8
Stampede	5	5	3
	All	5	3
Overall		30	34

ATEV also used the student t-test and a 90% significance level to test for a significant difference between the two groups for the number of non-detects or failures to detect (Table 4). Again, the analysis indicated there was a significant difference in the number of failures to detect between the two groups. More importantly, but not quantified, is the fact that failure to detect implies mission failure.

Finally, to insure that there were no differences within a group for the scenarios associated with the different EAs for each test (time to detect, number of false detections, and the number of failures to detects) the within group data was tested for homogeneity. The homogeneity tests confirmed that the data amongst EAs for each group for each test was not significantly different.

Issue 3. Situational Awareness and Risk Mitigation: Issues 3 and 4 were addressed qualitatively using a series of questions. Before the test subjects answered the questionnaires, the baseline group was exposed to the IRTSS capabilities and products. This was deemed necessary since some of the questions were of the form "In your opinion do you think an IRTSS pre-mission visualization capability would improve ....". Some of the questions posed to the test groups were:

- " Would this capability (IRTSS) improve pilot navigation, route selection and ingress/egress planning?"

- " Would you find a pre-mission FLIR look-back capability provides increased situational awareness useful for risk mitigation?"
- " Would pre-mission FLIR scenes assist to orient faster on an EA?"
- " Would a pre-mission FLIR planning tool help with crew coordination in BP operations?"

ATEC, based on the answers from both groups to the questions posed, concluded a "look-back capability in the opinion of the pilots increases situational awareness useful for risk mitigation" and "all pilots overwhelmingly feel that IRTSS would assist in situational awareness and risk mitigation".

Table 4. Number and Percentage of Failures to Detect the Target by Battle Position and Engagement Area

Engagement Area	Battle Position	IRTSS Number of Failure to Detects	IRTSS Percentage of Failure to Detects	Baseline Number of Failure to Detects	Baseline Percentage of Failure to Detects
Bravo	5	2	13	6	40
	6	2	13	5	33
	7	0	0	1	6.7
	All	4	8.9	12	26.7
Graham	3	0	0	0	0
	All	0	0	0	0
Shell	4	0	0	0	0
	5	0	0	0	0
	6	1	6.7	0	0
	All	1	2.2	0	0
Stampede	5	0	0	1	6.7
	All	0	0	1	6.7
Overall		5	4.2	13	10.8

Issue 4. Enhance Aviation Mission Planning: According to the findings of the ATEC analysis all pilots indicated that the delivery of pre-mission, synthetic FLIR scenes through AMPS would improve the IPB process planning and unit rehearsals.

CONCLUSION

IRTSS predicted synthetic FLIR scenes provide the warfighter with an operational capability not available today. It allows mission planners to directly and quantitatively account for weather when determining an optimum mission profile for both the tactical situation and the impact of environmental conditions on potential mission success. Enhanced aircrew situational awareness during mission execution is a second benefit associated with a pre-flight IRTSS capability. Pre-mission views of the EA IR clutter and the relative contrast between the target(s) and the immediate background facilitates long-range target detection and positive target identification (Bryant, 1998). IRTSS translates information dominance into readily assimilated situational awareness by fusing tactical intelligence with weather and depicting weather effects in a form that a non-meteorologist can easily understand and apply. The CEP has demonstrated that there is utility in providing Army Attack Aviators with pre-mission and pre-flight synthetic IR scenes generated from a physics-based model like IRTSS. If predicted FLIR capabilities are to enter the formal materiel acquisition process, the IRTSS technology is mature enough now to enter the life cycle at the system development and demonstration phase, thereby significantly reducing the time required for fielding (Milton and Williams, 2002).

REFERENCES

Balick. L. R., R. K. Scoggins, and L. E. Link 1981, Inclusion of a Simple Vegetation Layer in Terrain Temperature Models for Thermal IR Signature Prediction, *IEEE Transactions on Geoscience and remote Sensing*, GE-19(3):143-152.

Ballard, J. R. 1994 Grayling I Information Base Procedures for Generation of Synthetic Thermal Scenes: Final Report, SWOE Report 94-1, US Army Cold Regions Research Engineering Laboratory, Hanover, NH.

Berk, A., L.S. Bernstein. and D.C. Robertson, "MODTRAN: A Moderate Resolution Model for LOWTRAN 7', U.S. Air Force Phillips Laboratory, Geophysics Directorate, Hanscom AFB, MA 01731, Rept. No. GL-TR-89-0122, 30 April 1989.

Bryant, E. V., 1998, A Process Simulation Design to Assess Promising Technologies Relevant to F/A-18 Aircrew Target Recognition, Thesis: Masters of Science in Information Technology Management, September 1998, Advisor: William K. Krebs, Department of Operational Research, Associate Advisor: Terrance C. Brady. Department of Systems Management, Naval Postgraduate School, Monterey, CA 93943-5000.

Jordan, R. 1991 A one-dimensional temperature model for a snow cover: Technical Documentation for SNTHERM.89, USA Cold Regions Research and Engineering Laboratory, Special Report 91-16.

Koenig, G.G., Welsh, J.P., Wilson, J., Smart weapons operability enhancement synthetic scene generation process, SPIE—The International Society for Optical Engineering. Proceedings, 1995, Vol.2469, Targets and backgrounds: characterization and representation, Orlando, FL, Apr. 17-19, 1995. Edited by W.R. Watkins and D. Clement, p.254-265

Kress, R. M. 1992 Information Base Procedures for Generation of Synthetic Thermal Scenes: Final Report, SWOE Report 92-1, US Army Cold Regions Research Engineering Laboratory, Hanover, NH.

Milton, S. and R. Williams, 2002, "Mission Plannign and Rehearsal Tools for the Legacy, Interm, and Objective Forces," *Army AL&T*, May-June 2002, p 42-44.

Rynes, P., A. Curran, K. Johnson, D. Levanen, E. Martila, and T. Gonda, 2000, "Ground Vehicle Target Modeling with the Multi-Service Electro Optic Signature Code (MuSES)", *Eleventh Annual Ground Target Modeling & Validation Conference"*, Houghton, MI, 15-17 August 2000.

Seeley, G. and S. Luker, 1998, Infrared Target Scene Simulation Software, *Battlefield Atmospherics and Cloud Impacts on Military Operations [BACIMO] Conference*, AFRL Hanscom AFB, MA 12/1-12/3 1998, p. 340.

Siegel. S. and Castellan, N. J. Jr. 1988 *Nonparametric Statistics for the Behavioral Science (second edition)* New York: McGraw-Hill.

Smith, J.A., K.J. Ranson, D. Nguyen, L.K. Balick, L.E. Link, L. Fritchen, and B.A. Hutchison 1981 Thermal vegetation canopy model studies, *Remote Sensing of Environment,* 11:311-326.

Verhoef, W., and N.J.J. Bunnik 1975 A Model Study on the Relations Between Crop Characteristics and Canopy Spectral Reflectance, NIWARS publications No 33, 3 Kanaalweg Delft, The Netherlands, 89p.

Welsh, J.P., 1994, Smart Weapons Operability Enhancement (SWOE) Joint Test and Evaluation (JT&E) Program. Final Report, U.S. Army Cold Regions Research and Engineering Laboratory. SWOE report, Aug. 1994, No.94-10, Var. p., ADB-194 042.

Welsh, J.P., and L.E. Link Jr., 1995, Synthetic Scene Generation Process for Smart Weapons, Army RD&A, July-Aug, 1995, p33-36

Visualization Techniques For Four-Stokes Parameter Polarization

Grant Gerhart and Roy Matchko
US Army Tank-Automotive Research Development and Engineering Center
Warren, MI 48397
gerhartg@tacom.army.mil.

ABSTRACT

The human visual system can barely detect polarized light. This paper outlines a methodology for displaying and visualizing polarization profiles of natural, daylight scenes. A pseudo color scheme, based on the Poincare sphere, is developed for encoding the various polarization parameters in a scene. A mathematical model of the polarization state of skylight is given. A pictorial representation of the polarization state of the sky, using the Poincare colorization methodology, is presented. Using the colorization methodology, a visual correlation is given between the polarization states of light reflected from a target and the solar illumination producing it.

INTRODUCTION

The human visual system can barely detect polarized light. Some observers can determine, through sighting Haidinger's brush[1], that skylight is polarized. However, the human visual system is incapable of determining the complete polarization state of a light beam. Of course, we have instruments and methods that allow us to determine all the polarization parameters of a polarized beam of light. But how does one pictorially display all these parameters, especially if they vary spatially, as in a daylight scene? This paper is dedicated to our current work, which attempts to simply and conveniently solve this problem. The proposed solution involves a pseudo-color scheme that closely relates to the Poincaré sphere representation of polarization, developed by Henri Poincaré in 1892[2]. The surface of this sphere is commonly used to give a pictorial representation of all the possible polarization states of completely polarized light. Each point on the surface of the sphere corresponds to a unique polarization state. Right circular polarization is represented at the north pole and left circular polarization is represented at the south pole of the sphere. Linear polarization of all possible orientations is represented along the equator of the sphere. Elliptical polarization of all possible azimuth and ellipticity angles and handedness are represented at other points on the surface of the sphere.

Typically, only the surface of the Poincaré sphere is used to represent a polarization state[3]. Partially polarized light has been represented by a point moving along the surface of the sphere[4], the point spending more time in the neighborhood of a particular point. Our use of the Poincaré sphere is rather unconventional in that we use the inside of the Poincaré sphere as well as the surface of the sphere. The surface of the sphere corresponds to totally polarized light while the inside of the sphere corresponds to partial polarization, the center of the sphere corresponding to unpolarized light. This approach is in exact agreement with the one-to-one correspondence between the normalized Stokes parameters and the rectangular Cartesian coordinates of a point inside or on the surface of the Poincaré sphere.

Our main purpose in establishing a methodology for displaying and visualizing polarization states is to formulate a complete polarization profile of all objects in a natural, daylight scene. The recording of polarization information from a daylight scene can be accomplished through the use of a digital camera fitted with a filtering system that allows for the calculation of the four Stokes parameters[5].

This paper is a continuation of an extended effort to analyze ground combat vehicle signatures. Previous work[6-7] presented the results of a comprehensive study of the diurnal relationship between the slopes of target facets and edges, time of day and the azimuth angle of specular reflection of sunlight into a horizontal plane; the degree of polarization of reflected daylight from various paints and polarization profiles of diffuse and specular reflections from various paints and objects in daylight using a digital camera. In those previous studies, we showed that a complex index of refraction in the Fresnel reflection coefficients describes the degree of polarization for various paints under unpolarized light. Our goal then, as it is now, is to

understand the passive, visual signatures of targets in terms of phenomenological parameters such as angles of incidence and reflection, polarization angles, material properties, diurnal changes, vehicle geometry and shape, and scene content. Recent technological developments in the digital camera industry supply us with new tools to accomplish our goals and refresh previous interests.

Incident light on ground targets in daylight originates from two primary sources: sunlight and skylight. Sunlight is unpolarized, but diffuse skylight can be highly polarized due to Rayleigh scattering of sunlight[8-9]. Because the polarization state of skylight is different at each point on the celestial sphere, it is very difficult (if not impossible) to replicate this light source in the laboratory. One has no other choice but to study visual signatures related to daylight polarization parameters outdoors. Therefore, to study polarization signatures of targets in daylight, it is essential to have knowledge of the polarization state of the light originating from each and every point on the celestial sphere. A complete polarization profile of skylight need only include the degree of polarization and the polarization azimuth angle, since only linearly polarized light has ever been detected from skylight. Many authors give an equation for the degree of polarization due to Rayleigh scattering[10-12] in terms of the scattering angle. This paper gives an equation for both the degree of polarization and the polarization azimuth angle of skylight in terms of the position of the sun and the position of the point of observation on the celestial sphere in the horizon coordinate system. A pictorial representation of the polarization state of the sky, using the Poincare colorization methodology, is then presented. Using the colorization methodology, a visual correlation is given between the polarization states of light reflected from a target and the solar illumination producing it.

THE STOKES PARAMETERS AND THE POINCARÉ SPHERE

To determine the state of polarization corresponding to points in an image, three independent parameters must be determined for each pixel[13]. For example, the three independent parameters could be the amplitudes of the x and y components of the electric vector E_{0x} and E_{0y} and their phase difference, δ, along the optical axis. A prominent method to determine these three independent parameters, using measurable quantities, is the Stokes method[14]. This method involves measuring four intensities of a light beam. Each measurement corresponds to the intensity of the beam after it passes through each of four different filter system arrangements. The four Stokes parameters, sometimes called S_0, S_1, S_2 and S_3, are derived from these four measured intensities and form a four-element column vector in four-dimensional mathematical space. The Stokes parameters are applicable to a beam of light that is completely polarized, partially polarized or unpolarized; the beam may be monochromatic or polychromatic. The Stokes parameters for completely polarized light propagating along the +z-axis are

$$\begin{aligned} S_0 &= E_{0X}^2 + E_{0Y}^2 \\ S_1 &= E_{0X}^2 - E_{0Y}^2 \\ S_2 &= 2\, E_{0X}\, E_{0Y} \cos\delta \\ S_3 &= 2\, E_{0X}\, E_{0Y} \sin\delta \end{aligned} \qquad (1)$$

Since the primary purpose of this work is to determine the polarization states associated with pixels in an image of a scene, we are primarily interested in the relative values of the Stokes parameters. We obtain *normalized parameters* by dividing S_0, S_1, S_2 and S_3 by S_0. A normalized Stokes vector becomes $\{1, S_1/S_0, S_2/S_0, S_3/S_0\}$. The Stokes vector of a completely polarized beam is closely related to the Poincaré sphere representation illustrated in Figure 1. Every point on the surface of the sphere corresponds to a unique state of polarization of a plane monochromatic wave and all possible states of polarization have representation on the surface of the sphere. The parameter S_0 corresponds to the radius of the sphere and the

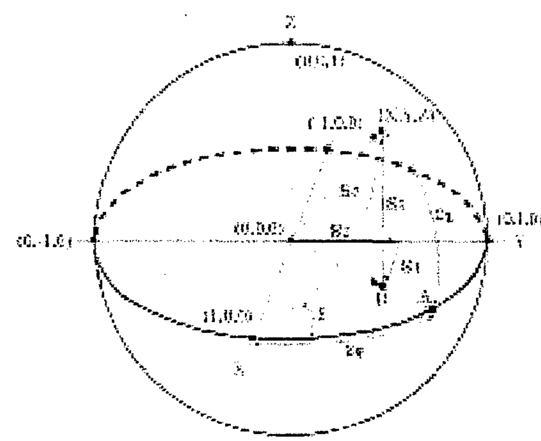


Figure 1. The relationship between the Stokes parameters and the Poncaré sphere.

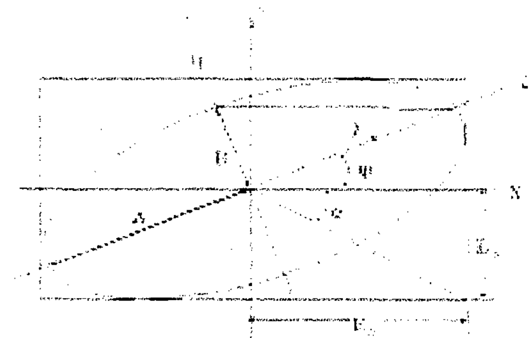


Figure 2. The polarization ellipse. The parameters ψ, χ are the polarization azimuth and ellipticity angles.

parameters S_1, S_2 and S_3 correspond to the rectangular Cartesian coordinates of a point (X,Y,Z) on the surface of the sphere. For partially polarized light the degree of polarization P is given by[15]

$$P = \frac{I_{Polarized}}{I_{Total}} = \frac{\sqrt{S_1^{\,2} + S_2^{\,2} + S_3^{\,2}}}{S_0} \tag{2}$$

where $I_{Polarized}$ is the intensity of the polarized component and I_{Total} is the total intensity. When the Stokes parameters are normalized, the degree of polarization P becomes the radius of a sphere with center coincident to the typical Poincaré sphere of unit radius (P = 1). The inside of the Poincaré sphere can be used to represent the polarized portion of partially polarized light (P < 1). The polarization forms on the surface of the inside spheres (P < 1) are exactly the same as the polarization forms on the surface of the sphere of unit radius. The benefit of this thought resides in the ability to colorize differently the same polarization form associated with a different degree of polarization. From the geometry of Figure 1 and Eq. (2) it can be shown that for normalized Stokes parameters and for any degree of polarization

$$S_1 = P\cos 2\psi \cos 2\chi,\; S_2 = P \sin 2\psi \cos 2\chi,\; S_3 = P \sin 2\chi \tag{3}$$

In fact, Eq. (3) represents the spherical coordinates for any point on the surface or inside of the Poincaré sphere where (x, y, z) = (S_1, S_2, S_3). The polarization azimuth angle, ψ, and the ellipticity angle, χ, are also defined in Figure 2. Using Eq. (3), ψ and χ can be determined from

$$\sin 2\chi = \frac{S_3}{\sqrt{S_1^{\,2} + S_2^{\,2} + S_3^{\,2}}} \qquad \text{and} \qquad \tan 2\psi = \frac{S_2}{S_1} \tag{4}$$

The parameter χ varies from +45° to -45°; it is positive for right-handed polarization forms and negative for left-handed polarization forms. The parameter ψ varies from 0° to 180°; it is 0° for horizontal polarization forms and 90° for vertical polarization forms.

ENCODING POLARIZATION PARAMETERS IN A DAYLIGHT SCENE

The Stokes parameters can easily be encoded in a daylight scene by assigning RGB values to the normalized values of S_1, S_2 and S_3 at each pixel site in the scene as follows:

$$R = \mathrm{int}[127.5\,(1 - S_1)],\; G = \mathrm{int}[127.5\,(1 - S_2)] \text{ and } B = \mathrm{int}[127.5\,(1 - S_3)] \tag{5}$$

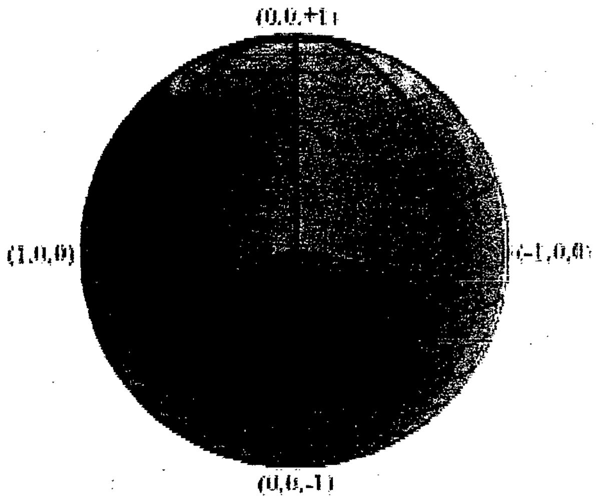
This pseudo-color scheme closely relates to the Poincaré sphere representation of partially polarized light where each Stokes vector maps into unique RGB values for 0*P*1. Unpolarized light ($S_1 = S_2 = S_3 = 0$) corresponds to middle gray (R = G = B = 127) at the center of the sphere (see Figure 3) while the inside of the sphere relates to partially polarized light ($0<P<1$). In this color scheme unpolarized or weakly polarized light is middle gray or highly unsaturated in the primary colors. The surface of the sphere consists of totally polarized light with P = 1. Figure 4 gives the appearance of some polarization forms, the pseudo-colors associated with them and their position on the Poincaré sphere. Figures 5 and 6 show the top and bottom views of the colorized Poincaré sphere, which relate to right and left-handed polarization forms respectfully.

The azimuth and ellipticity polarization angles are essential parameters in obtaining a complete polarization profile. One method of displaying these calculated parameters is to assign a different color to each specific angle. The equator of the Poincaré sphere plays a very special role in our pseudo-coloring scheme. It corresponds to linearly polarized light ($S_3 = 0$) and is also used to encode the polarization azimuth and ellipticity angles in a daylight scene (see Figure 7). Using $S_0 = 1$ and $\chi = 0$ in Eq. (3) and substituting into Eq. (5) yields

$$R = \mathrm{int}[127.5\,(1 - \cos 2\psi)],\; G = \mathrm{int}[127.5\,(1 - \sin 2\psi)] \text{ and } B = 127 \tag{6}$$

Substituting χ for ψ in Eq. (6) produces a color-mapping scheme for the χ-images.

$$R = \mathrm{int}[127.5\,(1 - \cos 2\chi)],\; G = \mathrm{int}[127.5\,(1 - \sin 2\chi)] \text{ and } B = 127 \tag{7}$$



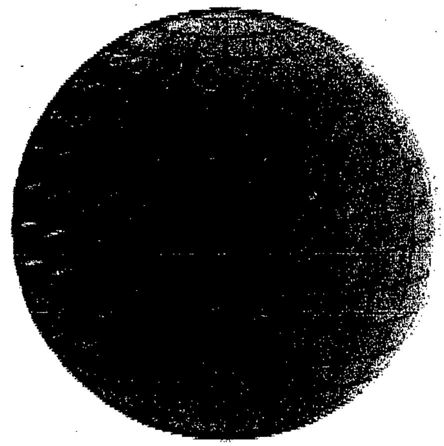


Figure 3. The colorized Poincaré sphere of unit radius. On the surface of the sphere P = 1, at the center P = 0, inside the sphere P < 1.

Figure 4. Pictorial representation of the polarization states corresponding to points on the surface of the colorized Poincaré sphere.

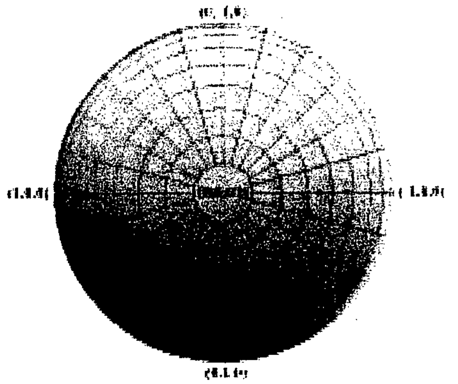


Figure 5. Top view of the colorized Poincaré sphere corresponding to right-handed polarization.

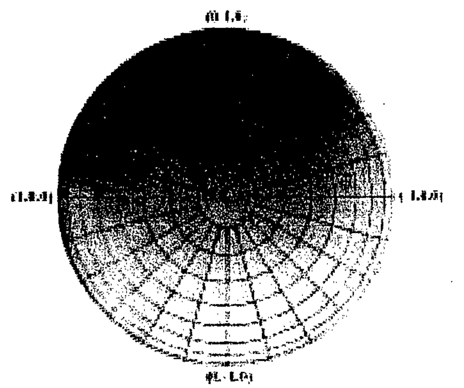


Figure 6. Bottom view of the colorized Poincaré sphere corresponding to left-handed polarization.

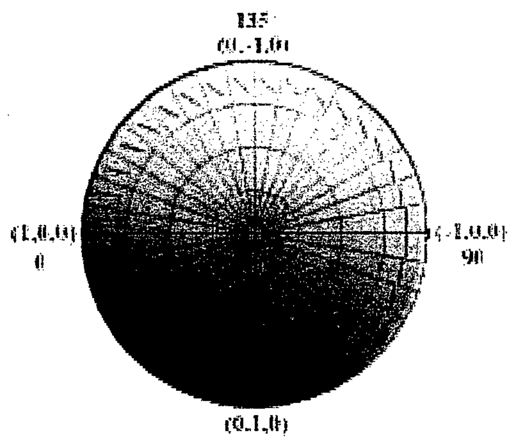


Figure 7. The Equatorial plane of the colorized Poincaré sphere.

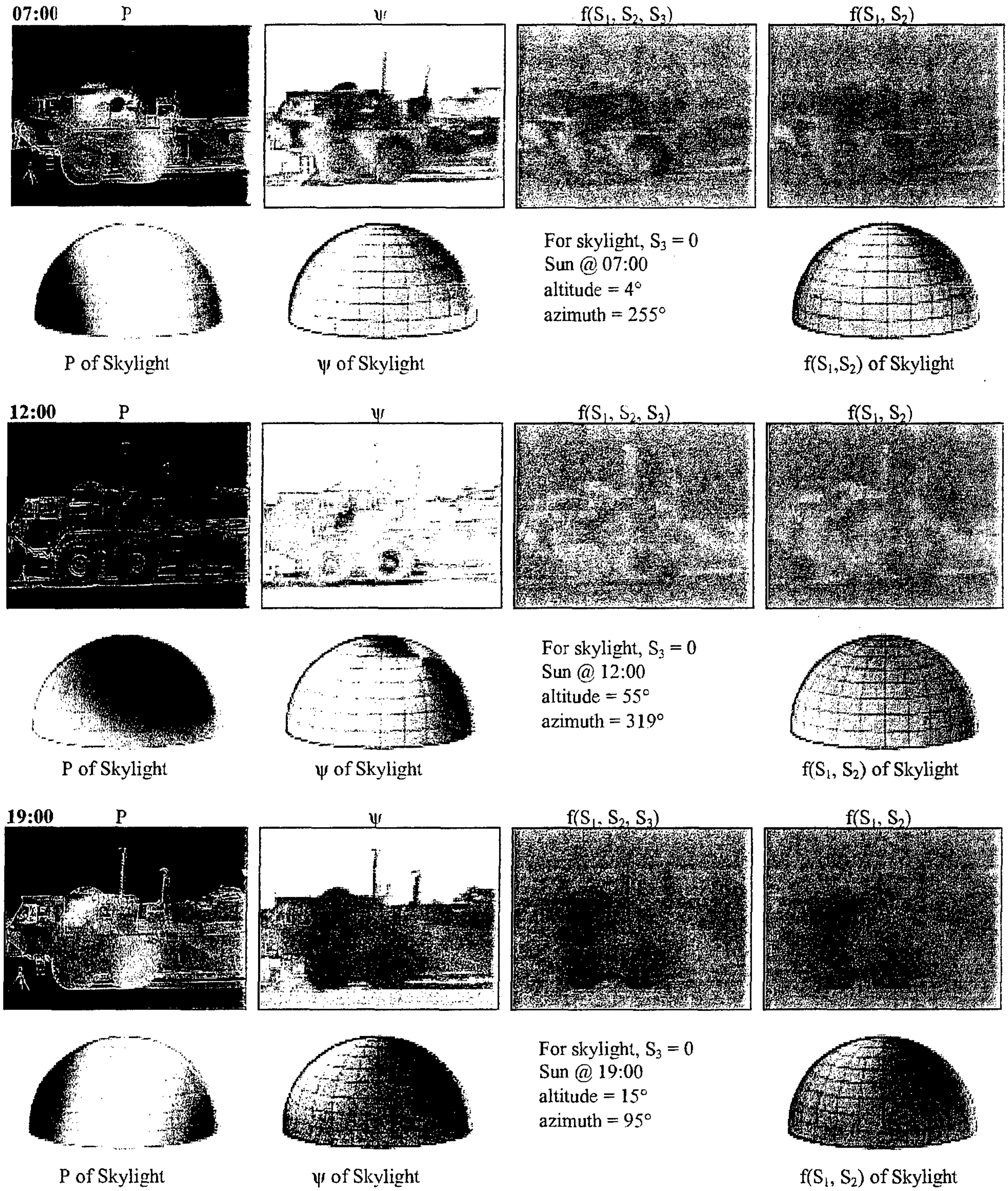


Figure 8. An example of the use of our encoding methodology to colorize the polarization parameters associated with a vehicle in daylight and polarized skylight. P = the degree of polarization, ψ = polarization azimuth angle, $f(S_1, S_2, S_3)$ and $f(S_1, S_2)$ = colorization according to Eq. (5). Azimuth convention: south = 0°, west = 90°, north = 180°, east = 270°.

There are certain cases where the Stokes parameter representation is ill conditioned. For example, as Eq. (4) shows, the azimuth angle is undefined when S_1 and S_2 are both zero (circularly polarized or unpolarized light). Also, the value of ψ and χ is undefined when $S_0 = 0$ (no light). Again, when the normalized value of S_3 approaches one (elliptically polarized light becomes nearly circular) the measured value of S_3 may exceed one due to signal to noise effects and produce an undefined χ-value ($\sin 2\chi > 1$). Since black (0,0.0), white (255,255,255) and yellow (255,255,0) are excluded from the encoding scheme of Eqs. (6-7), they are used for special conditions. Black is used when $S_0 = 0$, white for indeterminate values such as dividing by zero and for $P = 0$; yellow is used in the ψ-images whenever $\chi = \pm 45°$.

The degree of polarization, P, varies between 0 and 1. The simplest method of encoding this parameter in a daylight scene is to use the equation [pixel value] = 255 P. The black areas (pixel value = 0) in the resulting monochrome image will correspond to no polarization and the white areas (pixel value = 255) will correspond to light that is 100 percent polarized.

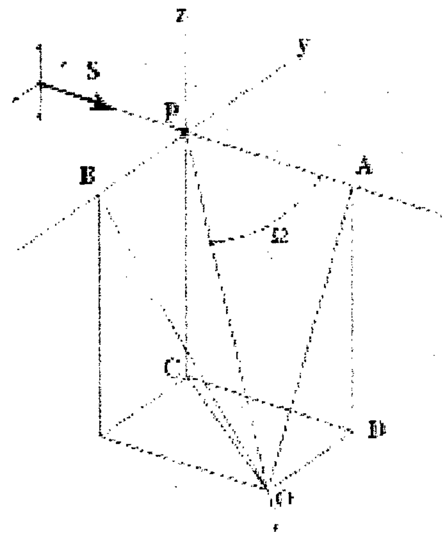


Figure 9. The scattering angle Ω.

Figure 8 gives an example of the use of our encoding methodology to colorize the polarization parameters associated with a vehicle in daylight. The images relating to Figure 8 were captured during clear skies, near sunrise, noon and sunset on April 26, 2000, at north latitude 42.5 degrees and west longitude 83 degrees. A normal to a vertical vehicle panel, which faces the digital camera, points south; the camera view is toward the north. A comprehensive diurnal study of this vehicle involved acquiring enough images to calculate the Stokes parameters every 30 minutes from sunrise to sunset. Several derived images from this study are shown in Figure 8. The notations $f(S_1, S_2, S_3)$ and $f(S_1, S_2)$ refer to the encoding scheme given in Eq. (5). The notation $f(S_1, S_2)$ implies $S_3 = 0$, which is the case for the polarization of skylight. Hence, any difference that occurs between the $f(S_1, S_2, S_3)$ and $f(S_1, S_2)$ images is due to elliptically polarized light.

The skylight images below the vehicle images in Figure 8 are pictorial representations of the polarized sky in the southern semi-hemispherical region of the celestial sphere. An Excel spreadsheet was used to create the gridlines. MatLab scripts were written to colorize the areas between the grid lines according to the encoding scheme given in Eqs. (5-6). The skylight polarization parameters were encoded into the images as they would appear looking toward the earth, as from a satellite above the earth. The values for the observed degree of polarization of skylight, along a line of sight, were obtained from[9-11]

$$\mathbf{P} = \frac{\sin^2\Omega}{1+\cos^2\Omega} \tag{8}$$

where Ω is the scattering angle and is defined in Figure 9. Figure 9 shows a sun's ray S incident on particle P and the scattered light observed along the line of sight PO. The scattering angle, expressed in the horizon coordinate system, is given by

$$\cos\Omega = -\cos\alpha\cos\theta\cos(A-\varphi) - \sin\alpha\sin\theta \tag{9}$$

where α = the altitude of the sun and A = the azimuth of the sun, θ = the altitude of the point of observation and φ = the azimuth of the point of observation (south = 0° and lies in the + y direction, west = 90° and lies in the + x direction). According to Eqs. (8), if $\Omega = 90°$, scattered sunlight is 100 % linearly polarized at right angles to the observation plane. For angles $0° < \Omega < 90°$ and $90° < \Omega < 180°$ scattered sunlight is partially polarized. The maximum intensity of scattered sunlight occurs when $\Omega = 0°$ or $180°$, in the forward and backward direction. Figure 10 shows the geometry for finding the polarization azimuth angle produced by scattered sunlight. R is a unit position vector pointing toward the sun and Q is a unit position vector in the direction of observation. Since a tangent **T** to the polarization circle passing through the observation point is perpendicular to both **R** and **Q**, it can be calculated from the vector cross product **T** = **Q X R**. A vector **N**, perpendicular to the plane containing the z-axis and **Q**, is obtained from the vector cross product **N** = **Q X k**, where **k** is a unit

vector along the z-axis. The angle between T and N is the polarization azimuth angle ψ observed at the observation point $P_o(x_o,y_o,z_o)$ along the direction of Q and is obtained from a vector dot product. The value of ψ can be obtained from

$$\cos\psi = \frac{\sin\alpha\,\cos\theta - \sin\theta\,\cos\alpha\,\cos(\varphi - A)}{\pm\sqrt{\begin{aligned}&(\cos\theta\,\sin\alpha\,\cos\varphi - \sin\theta\,\cos\alpha\,\cos A)^2 + \\ &(\sin\theta\,\cos\alpha\,\sin A - \cos\theta\,\sin\alpha\,\sin\varphi)^2 + \\ &\cos^2\theta\,\cos^2\alpha\,\sin^2(\varphi - A)\end{aligned}}} \tag{10}$$

The inputs are the altitude α and the azimuth A of the sun and the altitude θ and the azimuth φ of a point of observation. Azimuths are measured from due south toward the west and around the horizon 360 degrees (south = 0°, west = 90°, north = 180°, east = 270°).

If $\psi < 90°$, a line L tangent to a skylight polarization circle and passing through the point $P_o(x_o,y_o,z_o)$, will pierce the x-y plane to the left of the vertical plane P, as shown in Figure 11; otherwise, if $\psi > 90°$, it will pierce the x-y plane to the right of plane P.

This simple picture of the polarization produced by the scattering of sunlight is incomplete for the following main reasons: (1) multiple scattering, (2) molecular anisotropy and (3) size of particles. However, Eqs. (8-10) are quite adequate in describing the polarization parameters associated with large regions of the sky.

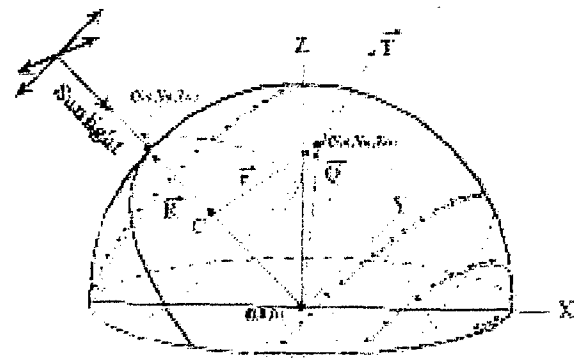


Figure 10. Geometry for finding the polarization azimuth angle produced by scattered sunlight.

ANALYSIS OF ENCODED POLARIZATION PARAMETERS IN A DAYLIGHT SCENE.

Many factors determine the nature of the reflected light from objects in a daylight scene that are recorded by a digital camera. The nature of the incident light, the nature and geometry of the material the incident light interacts with, the object-light source geometry and the spatial geometry between the camera, the object and the light source. The primary sources of light in a daylight scene are direct sunlight and skylight. Sunlight is unpolarized and skylight has polarization characteristics described by Eqs. (8-10). Commercial computer software is available to determine the position of the sun at any location and time on earth[16]. Since the view of the digital camera (the one used to acquire the images in Figure 8) was to the north and its optical axis was horizontal, the light that entered the digital camera was approximately horizontal and directed toward the south.

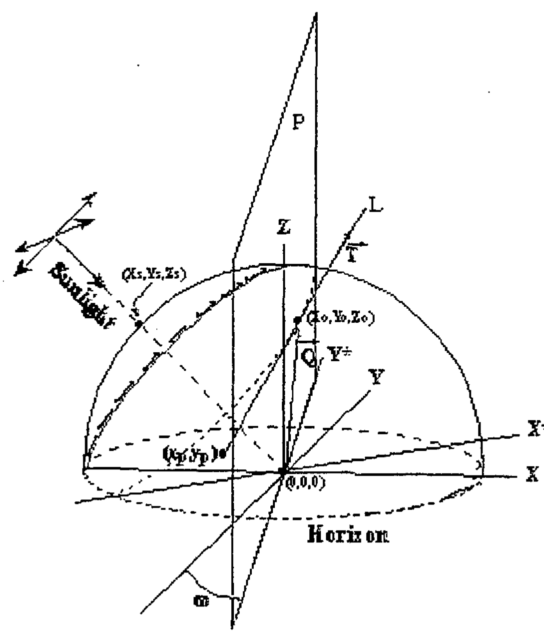


Figure 11. Geometry for finding the piercing point (x_p,y_p) in the x-y plane.

An examination of the encoding of polarization parameters in the daylight scenes of Figure 8 reveals a feature common to all images; namely, the color of the target affects the polarization state of the reflected light. The door panel of the vehicle, which lies in a vertical plane, is a good example of this phenomenon. The reflected light from those parts of the door that have the same color have approximately the same degree of polarization, azimuth angle and ellipticity angle. But, different colors on the door panel have different P, ψ and χ-values.

The P-images of Figure 8 (degree of polarization) at 07:00 and 19:00 show much less contrast and higher values of P than does the image at 12:00. At 07:00 and 19:00, the skylight images (below the vehicle images in Figure 8) show that highly polarized skylight exists at all altitudes in the southern sky, i.e., along the meridian and directly opposite the vehicle. Figure 12 shows that both edges and vertical panels can reflect highly polarized light into the camera. Since the value of P is large for both of these regions, there is little contrast between them in the P-images. The sun has yet to cross the vertical plane of the vehicle at 07:00 and is in the western sky, near the horizon, at 19:00. Neither of these positions of the sun can contribute much to producing highly polarized reflected light from the vehicle. At 12:00, the altitude and azimuth of the sun are 55° and 319° respectfully. Highly polarized skylight is near the east and western portions of the visible sky. Hence, skylight contributes very little to the degree of polarization of reflected light from vertical panels. As Figure 13 illustrates, only sunlight that reflects specularly from edges can enter the camera at this time; specular reflections from vertical panels are toward the ground. However, sunlight that reflects diffusely from both edges and vertical panels can enter the camera. In-house studies, with surfaces coated with the same material as the vehicle in Figure 8, show that the highest degree of polarization (0.55) occurs for obtuse angles between the incident light and the light reflected from the surface. Since obtuse diffuse angles of reflection for vertical panels are directed toward the ground, the edges will show a larger degree of polarization than vertical panels. This results in more contrast between the vertical panels and edges of the vehicle in the P-images.

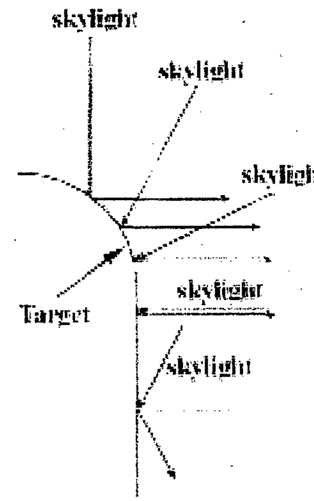


Figure 12. At 07:00 and 19:00, highly polarized skylight along the meridian is incident on the target.

In-house studies, of colored panels painted with the same material as the vehicle shown in Figure 8, show that diffuse reflections from vertical panels, <u>for all angles of incidence</u>, produced by *<u>unpolarized</u>* incident light (analogous to sunlight), have polarization azimuth angles of 90°. This implies that the electric field vectors of the reflected light are parallel to the surface. In-house studies also show that diffuse reflections, produced by *<u>polarized</u>* incident light (analogous to skylight), have polarization azimuth angles that are color dependent. One of the colors showed that diffuse reflections from vertical panels, produced by *<u>polarized</u>* incident light, <u>for all angles of incidence</u>, have polarization azimuth angles that are the mirror images of the azimuth angles of the incident light. In other words, with regards to reflection azimuth angles produced by polarized incident light, this particular colored target performs like a mirror. In-house studies also show that the largest degree of polarization of reflected light that originates from *<u>unpolarized</u>* incident light is 50% as compared to nearly 100% for *<u>polarized</u>* incident light. Hence, skylight will tend to dominate over sunlight, with regards to measured polarization azimuth angles, ψ, of reflected light from the vehicle. This behavior of ψ is seen in the daylight scenes of Figure 8. A comparison of the ψ-images of the vehicle and the ψ-images of skylight, for the same time of day, shows that large regions of the vehicle have reflected azimuth angles that are approximately the mirror image of the skylight azimuth angles near the meridian and near the horizon, i.e., in the direction of the digital camera. Since ψ is determined from the ratio of S_1 and S_2, as given by Eq. (4), the $f(S_1, S_2)$-images of the vehicle in Figure 8, as determined from Eq. (5), have similar colorized regions as the $f(S_1, S_2)$-images of skylight near the meridian and near the horizon, i.e., in the direction of the digital camera.

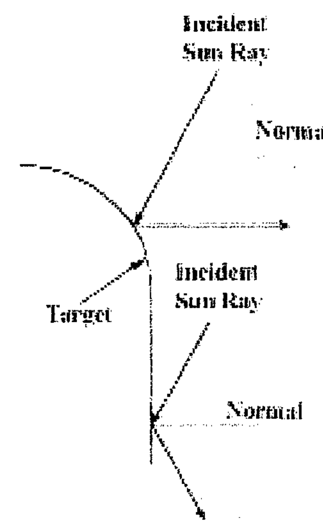


Figure 13. At 12:00, only specular reflections of sunlight from edges can enter the camera; specular reflections from vertical panels are toward the ground.

In-house studies, relating to the ellipticity angle of light reflected from panels painted with the same material as the vehicle shown in Figure 14, show the following:

1. The ellipticity angle shows some dependence on the color of the surface coating of the target.
2. For incident unpolarized light and polarized light and for diffuse reflections, the ellipticity angle χ increases with decreasing values of the degree of polarization P.

3. For diffuse reflections (scattering) and incident unpolarized light, the largest values of χ and the smallest values of P occur for backward reflections (back scattering).
4. For diffuse reflections and polarized incident light. the largest values of χ and the smallest values of P occur for incident polarization azimuth angles between 20° and 30° and for diffuse angles of incidence, I, around 100°.

In the diurnal daylight studies shown in Figure 14, the largest values of χ and the smallest values of P occur in the 12:00 image. During this time, the angle between the optical axis of the camera and the position along the horizon of the highest skylight degree of polarization was approximately 90° (east and west). Also, the skylight polarization azimuth angle during these times, near the horizon and in the east and west, was approximately 20° and 30°; in excellent agreement with the results of in-house experiments. The 07:00 images, which correspond to the sun to the east of the meridian, show primarily right-handed polarization, where as the 19:00 image, which corresponds to the sun to the west of the meridian, shows primarily left-handed polarization.

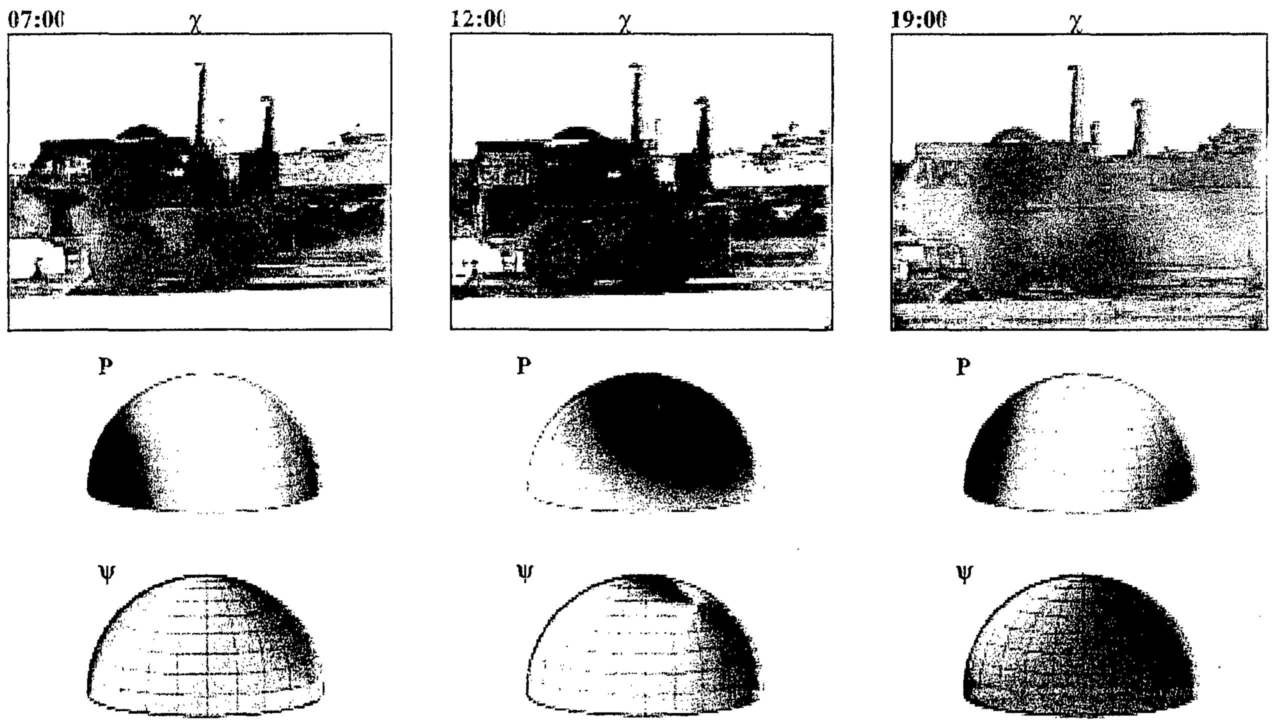


Figure 14. Encoding of the polarization ellipticity angle χ associated with a vehicle in daylight, the degree of polarization P of skylight and the polarization azimuth angle of skylight ψ.

CONCLUSION

This paper has outlined a methodology for displaying and visualizing polarization profiles of natural, daylight scenes. A pseudo color scheme, based on the Poincaré sphere, was developed for encoding various polarization parameters. Examples were given of encoding the degree of polarization and the polarization azimuth and ellipticity angles in daylight scenes at different times of the day. The scenes contained targets with surface coatings that were studied in the laboratory. In-house results, obtained from a luminance meter, were found to be in close agreement with results obtained out-doors, using a digital camera.

A mathematical model of the polarization state of skylight was given in terms of the horizon coordinate system. A pictorial representation of the polarization state of the sky was presented that used the Poincaré colorization methodology and

corresponded to the times of the sample daylight scenes. Using the Poincaré colorization methodology, a visual correlation was given between the polarization states of light reflected from a target and the solar illumination producing it.

REFERENCES

1. M. Minnaert, *The Nature of Light and Color in the Open Air*, Dover Publications Inc. (1954).
2. H. Poincaré, *Théorie Mathématique de la Lumière II.* Paris, Geores Carré, editeur (1892), pp. 275-306.
3. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, North-Holland, Amsterdam, 59 (1992).
4. Ibid., p. 54.
5. G. Gerhart and R. Matchko, "Method of determining polarization profiles for polychromatic sources," United States Patent Office, Washington D.C., Patent Number 5,734,473, Col. 4, (1998).
6. R. Matchko and G. Gerhart, "Luminance, Contrast and Polarization of White Light Reflected from Ground Combat Vehicles", No. 13568, TACOM Research Development and Engineering Center, Warren, MI (1992).
7. R. Matchko and G. Gerhart, "Reflection and Polarization of White Light from Ground Vehicles", ADB271426, TACOM Research Development and Engineering Center, Warren, MI (1998).
8. F. Jenkins and H. White, *Fundamentals of Optics*, 3rd ed., McGraw Hill, New York, pp. 505-507 (1957).
9. E. Hecht, *Optics*, 2nd ed., Addison-Wesley, Reading, Massachusetts, pp.292-296 (1990).
10. W. J. Humphreys, *Physics of the Air*, chs. 7-8, Dover Publications (1964).
11. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, sec. 5.1, John Wiley & Sons, New York (1983).
12. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, New York, pp. 652-656 (1993).
13. Ibid., pp. 24-31.
14. G.G Stokes, "On the composition and resolution of streams of polarized light from different sources," Trans. Cambridge Phil. Soc. 9, 399 (1852); *Mathematical and physical papers*, vol. 3 (Cambridge University Press, Cambridge, England, 1901), p. 233
15. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, New York, pp. 555 (1993).
16. Software Bisque, The Sky Astronomy Software, Golden, CO, www.bisque.com.

A Mobile Decision Aid for Determining Detection Probabilities for Acoustic Targets

David Sauter, David Marlin
Army Research Laboratory
White Sands Missile Range, NM 88002 USA

ABSTRACT

The Army Research Laboratory (ARL) has developed a physics based Acoustic Battlefield Aid (ABFA) for acoustic sources and sensors. This application computes numerous output parameters (e.g., probability of detection, transmission loss, signal-to-noise ratio, etc) for a number of acoustic sources and sensors (if required). Probability of detection (POD) output would seem to be particularly of interest to lower echelon units, e.g., in determining how close enemy tanks could approach before being heard (by either the human ear or some other acoustic sensor). As a result, a prototype application has been developed that allows the POD for a user specified target/sensor pairing given the direction and distance between the two. Commercial wireless communications are used to link the mobile device (a Compaq 3650 personal digital assistant) to a remote laptop server. The server consists of a set of dynamic linked libraries written in C that contain several functions that access an acoustic propagation table. This table is a function of the environmental conditions, and for the prototype, is simply a static table. Eventual dynamic creation of the table from an existing gridded prognostic meteorological database (residing on the Army's tactical command and control system for weather) is anticipated. Tactical wireless comms would also replace the short-range commercial comms.

INTRODUCTION

While military ground targets have traditionally been detected via their visible and infrared emissions, detection of their acoustic signatures can also be used for target acquisition. For example, the Brilliant Anti-Armor Technology Submunition (BAT) uses an acoustic sensor to seek out armor targets, and an infrared sensor to engage the vehicles. As with visible and infrared signatures, the atmosphere can attenuate the transmission of the acoustic signal. Realizing the requirement for providing information regarding the acoustic source propagation through the environment and its effects on the ability of sensors to detect the signal, ARL researchers have developed an Acoustic Battlefield Aid (ABFA) [1]. ABFA is a physics based model that can provide numerous outputs regarding the acoustic signal transmission through the atmosphere. Some of the outputs available include probability of detection, direction finding accuracy and signal-to-noise ratio. Inputs include the meteorological conditions, underlying terrain type (forest, gravel, sand, etc) the acoustic source (target – T62 at 19 mph, generic tank, stationary HMMWV, etc) and receiver (sensor – human listener, generic single microphone, microphone array, etc). It is currently available on Windows based platforms although an effort is underway to provide the ABFA functionality on the Army's fielded tactical command and control system for weather, the Integrated Meteorological System (IMETS). Figure 1 is a capture of a recent ABFA graphical user interface and display showing the detection probability for a generic tank by a microphone array. The interface also shows a pull down source menu that can be used to select a specific acoustic source. Weather data can currently be entered manually, downloaded from the internet or the user can choose pre-defined meteorological cases, e.g., "clear night, moderate wind". When integrated into the IMETS, the weather data can be automatically retrieved for the specified location by accessing a gridded database containing the output from a prognostic fine scale weather model.

As can be seen from Figure 1, the ABFA GUI is rather involved and contains numerous menu items and options that must be set by the user. This is beneficial in terms of providing a powerful and highly configurable modeling environment, however, for lower echelon users the computing platform that they have access to may not be amenable to this level of detail. In this case, when the majority of the meteorological and terrain data can be maintained on a server platform (e.g., an IMETS at battalion level), a simple GUI on a handheld computing device may suffice. This GUI would provide the user with enough flexibility to specify the basic input parameters (e.g., source and sensor as well as azimuth and range between them) that can then be sent to the server (over wireless comms, for instance) for computation of probability of detection (or some other output parameter). Thus, critical quantitative detection related information could be provided to lower echelon users on the local device. The remainder of this paper discusses such a prototype application that has been developed and

demonstrated by ARL.

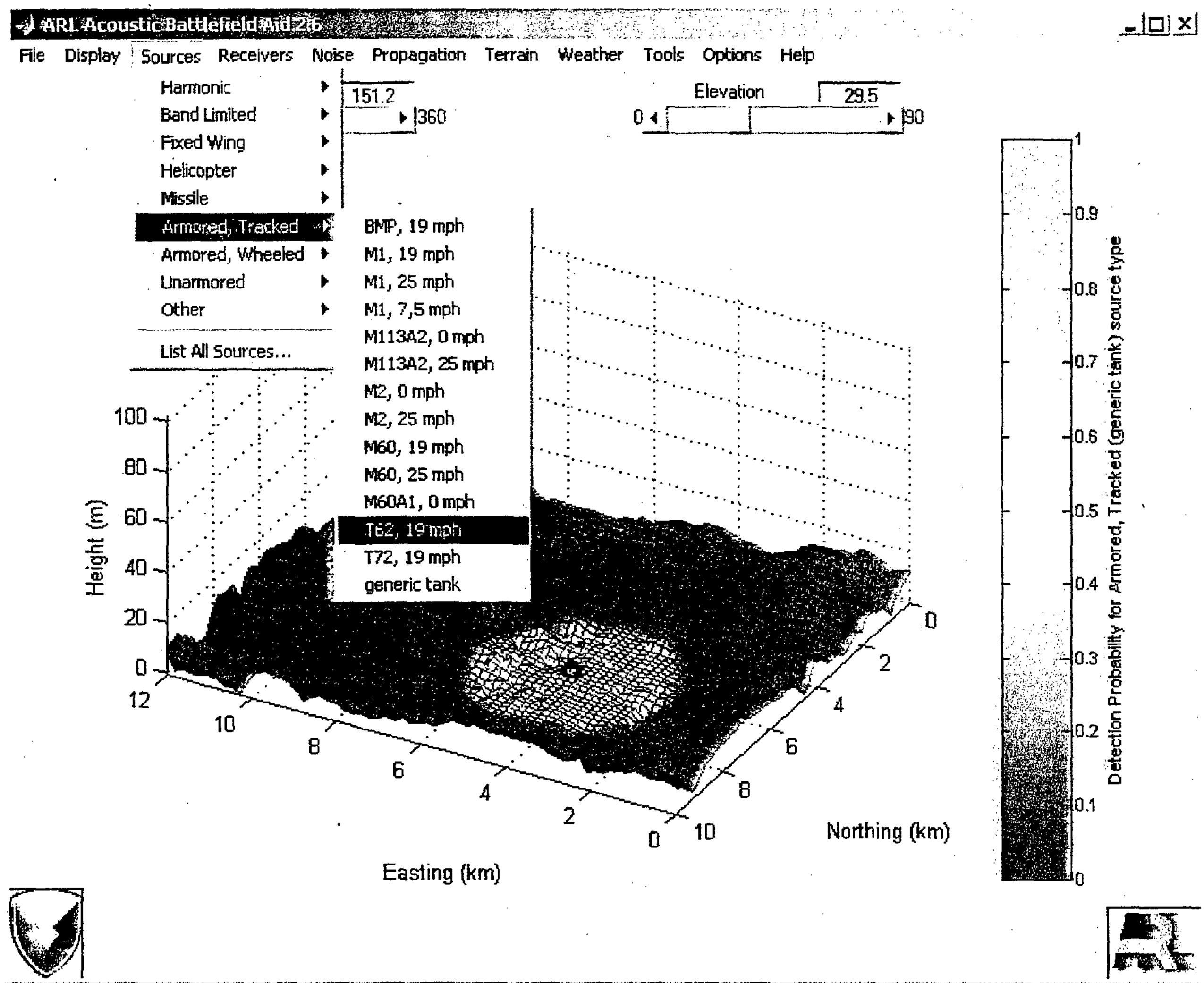


Figure 1 - ABFA Graphical User Interface

REQUIREMENT

The Army's Objective Force will emphasize highly mobile and independent fighting units. To enable the Objective Force, information and situational dominance will be required at all echelons. This will require providing critical information to the lower echelons via small computing devices (e.g., handheld computers) over wireless communications. Although an IMETS will be fielded at battalion level via a light version on an Intel based hardware platform, there are currently not any plans to field an IMETS at echelons below this. In an effort to address this requirement, the Army is developing prototype hardware and software environments to provide a lower echelon computing capability on handheld or mobile computing devices. For example, there are ongoing efforts at the Dismounted Battlespace Battle Lab (Ft. Benning, GA), the Land Warrior Program (Ft. Belvoir, VA) and the Objective Force Warrior (Natick, MA) related to mobile computing applications. ARL is involved in developing prototype applications specifically related to weather and weather impacts on weapon systems and military operations, of which acoustic propagation is one.

Even though there have been numerous technological advances in the mobile computing environment, it is not yet feasible to make all of the required computationally intensive calculations on the mobile device, thus a client-server environment was

chosen to develop and demonstrate our prototype Mobile Acoustic BattleField Aid (MABFA). The specifics of both the server and client are discussed briefly in the following section.

CLIENT-SERVER ENVIRONMENT

Surrogate Server

Detailed acoustic propagation information is computed as a function of the environmental conditions (wind velocity, temperature profile, etc) and stored in tables on a surrogate IMETS server (a Pentium 450 MHz Windows 2000 based laptop computer in this case). These tables are computed from algorithms written in the C programming language and derived from the ABFA code. Currently the tables are static for demonstration purposes. However, they will eventually be derived from the gridded meteorological database that is resident on the IMETS. This database is both temporally and spatially dependent and provides relatively high fidelity (10 km in the horizontal) prognostic (48 hours) data from which the propagation tables can be computed. These tables will be dynamically updated twice daily when the gridded meteorological database is refreshed. Depending on the compute time for the tables, they will either be computed a priori and stored for each grid point or they will be computed on the fly for a specific grid point that is nearest the location of the mobile user requesting acoustic propagation information. The IMETS gridded database is typically valid for a 500x500 km domain, thus there are 2601 grid points (51x51) in the horizontal.

Client Environment

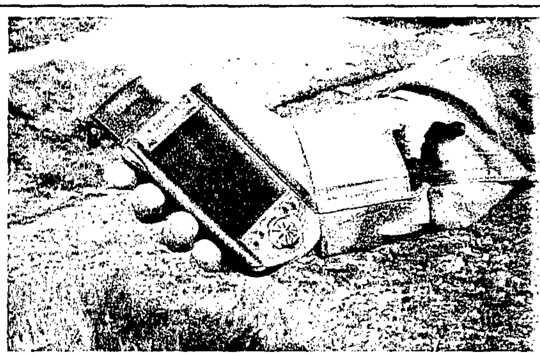


Figure 2 – Compaq 3650 with 802.11 PC Card

The client device is a Compaq personal digital assistant, specifically model 3650 (see figure 2). Although the device weighs just over 6 ounces and has approximate dimensions of 5" x 3.3" x 0.6", it is a fairly capable platform with a 206 MHz processor, color display, integrated microphone and speaker, and 32 Mb of memory (newer versions have 64 Mb). Enhancing the device's capability is the PCMCIA[1] expansion pack option. This allows either one or two PC cards to be integrated with the mobile platform. Current cards supported include 802.11 based wireless comms[2], GPS, additional memory via compact flash cards or micro drives (up to 10 Gigabytes!), and VGA output to an external display. Of primary interest is the wireless capability which is currently being used to transmit and receive information from the surrogate IMETS server over distances of up to 1000 feet via a small antenna plugged into an external wireless access point connected to the server. The wireless comms utilized for the prototype application could be replaced with tactical comms for fielding, e.g., possibly the emerging Joint Tactical Radio System (JTRS) which will be able to transmit/receive both voice and data. The GPS capability can be integrated with the acoustics application to pass the geographic location of the mobile device and then query the propagation tables nearest to that point. For more detailed information regarding the mobile computing device and the software environment utilized see a recent journal article by Sauter and Torres [2].

The client application is written in Java for portability and potential transition to a fielded system. The code as written and compiled on a Windows 2000 operating system machine runs *without recompilation* on the Compaq device as well as Windows 95 and Solaris based platforms. Java's remote method invocation (rmi), a set of protocols to allow Java objects to communicate with each other remotely, is used for the client-server interaction. Since it was stated above that the server software is written in the C programming language, another interface is required on the server end to allow the Java client to access the propagation tables. This interface is the Java Native Interface (JNI) and allows Java objects to interact with C functions. RMI and JNI are both relatively straightforward to work with and did not pose any challenges to the implementation of the acoustic propagation mobile application.

[1] Personal Computer Memory Card International Association, an organization of some 500 companies that has developed a standard for small, credit card-sized devices, called PC Cards (from www.webopedia.com).

[2] 802.11 refers to a family of specifications developed by the Institute of Electrical and Electronics Engineers for wireless LAN technology. 802.11 specifies an over-the-air interface between a wireless client and a base station or between two wireless clients (from www.webopedia.com).

MOBILE ACOUSTIC BATTLEFIELD AID

Once the client and server environments were chosen, it was relatively easy to develop the client application. A graphical user interface (GUI) was developed Java on a laptop computer and then copied to the Compaq mobile device (although Java code can be run on the mobile device, there is not a development environment on it, thus code must be compiled elsewhere and transferred). While the full Acoustic Battlefield Aid (ABFA) application has numerous inputs and possible outputs, it was felt that for a mobile application, the interface and number of output parameters should be kept simple. Thus, an application that would provide the probability of detection of an acoustic source by a sensor given the range and azimuth between the two was coded. Figure 3 displays the actual mobile ABFA GUI as it appears on the mobile device. Pull down menus allow the user to choose one acoustic source and one sensor. The Java programming environment for the mobile device is rather versatile and allows for error trapping of invalid inputs (e.g., a character entry for a numeric input field or an azimuth value > 360 degrees). In turn, the user cannot submit the request for the server computations until all inputs contain valid entries. A text area is used to display pertinent information to the user as is displayed in the GUI in the figure. Sound can also be added to the client application to alert the user of invalid entries, problems with the server or successful retrieval and display of information from the server.

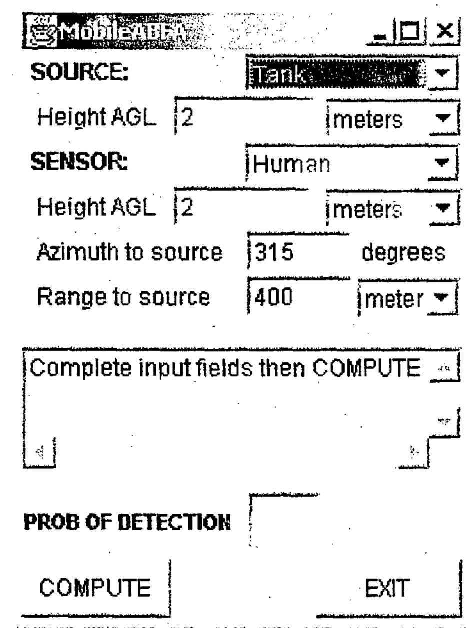


Figure 3 – Mobile ABFA GUI

For the prototype application the source and sensor are actually fixed on the server. Once the user has entered the range and azimuth and clicks the "COMPUTE" button shown in the lower left portion of the GUI, the information is transmitted to the surrogate IMETS server over the wireless comms. The computations are then made using the static propagation tables and the probability of detection information is sent back to the client and displayed. The time from "COMPUTE" to the display of the information on the mobile device is typically a few seconds. Depending on the meteorological scenario, the probability of detection can vary greatly depending on not only the range between the source and sensor but also on the azimuth. Acoustic signals can be transmitted longer distances along the downwind direction than in the upwind direction. Thus, if the azimuth is aligned in the same direction as the wind direction and the sensor is downwind of the source, the probability of detection can be much larger than would otherwise be expected.

The Mobile ABFA application should provide useful information to the mobile user regarding detection probabilities/distances not only for enemy vehicles, but also for estimating possible avenues of approach for friendly vehicles on enemy positions. For example, because of terrain or other features there may exist two potential avenues of approach upon an enemy location. If one is more vulnerable than the other to acoustic detection by enemy forces, a decision could be made to use the approach with the lower probability of acoustic detection. Depending on mobile user requirements, it would be relatively simple to add output parameters to the Mobile ABFA application other than the probability of detection. As was stated earlier, there are multiple parameters that are available from the ABFA program. Thus, if the user wanted to see transmission loss that value could be computed on the server and transmitted back to the mobile application for display.

FUTURE CAPABILITIES

2-D/3-D Displays

The full ABFA application relies heavily on 2-D and 3-D displays to display spatially dependent output information to the user (see Figure 1). It should be possible in future versions of the Mobile ABFA to display some of this information for the mobile user. This would be possible via two separate methods. Both would involve the computation of the spatially dependent information on the server. In the first method, the actual 2-D or 3-D display would also be created on the server and saved as a graphics image in the Joint Photographic Experts Group format (jpeg or jpg for short). As the mobile device has an image viewer to display jpg images, the file could be transmitted to the mobile user over the wireless comms and then loaded and displayed. The drawback to this approach is that the bandwidth for the existing wireless comms is relatively small (1 megabit per second in theory, but in practice typically significantly less) and some of these images could be relatively

large (over 1 megabyte). Thus, if there are several mobile users requesting multiple images the transmit times could be long. In addition, in a fielded version there would be competing applications requiring use of the wireless network, effectively decreasing the bandwidth even more. A second option would involve transmitting the spatially dependent information to the mobile user and then composing the image on the mobile device. Commercial mapping software solutions for the Compaq 3650 do exist (e.g., ESRI's ArcPad), so in theory, this information should be able to be rendered on the mobile device. The amount of information required to be transmitted would vary with the domain of interest and the resolution of the grid points but should result in a lesser amount of data transferred between server and client. Benchmarking tests would need to be performed to determine the tradeoffs/benefits of each method.

Target Detection via non-Acoustic Signals

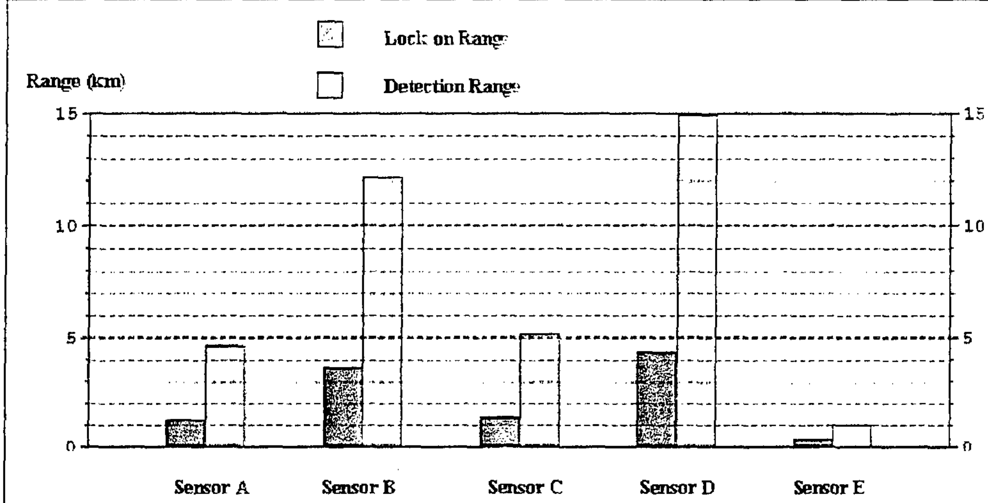


Figure 4 – TAWS Bar Chart of Lock On and Detection Ranges

The Target Acquisition Weapons Software (TAWS [3]) provides lock on and/or detection ranges for a wide variety of electro-optical sensors and military targets. Although this software is currently written in Fortran, it should be possible to develop a mobile client in Java that would display TAWS output parameters. This would compliment the Mobile ABFA application and allow the lower echelon user to obtain valuable information regarding target acquisition in the electro-optical wavelength spectrum. Currently a version of TAWS is implemented in the IMETS and has non-map graphics displays to represent the lock on and detection ranges as bar charts (figure 4). It is possible that at some point these bar charts could be displayed on the mobile device, but to start with it may be simpler to modify the existing Mobile ABFA GUI such that lock on and detection ranges can be displayed as numeric values in the current Probability of Detection text box. Work is being finalized on a TAWS map overlay capability in IMETS that will overlay the lock on and detection ranges on a map background over a 360-degree azimuth. This overlay will also incorporate the effects of terrain masking and will color code those areas in a separate color. The discussion in the section above related to providing a 2-D or 3-D display on the mobile device is pertinent to the overlay capability as well. At some point, it may also be possible to animate these graphics over time to allow the user to visualize the temporal variation in the transmission of the acoustic and electro-optical signals.

SUMMARY

Advances in computer hardware and software technology have made it possible to develop and demonstrate a prototype Mobile Acoustic BattleField Aid on a handheld computer. While the current GUI and outputs are rather simplistic, it is anticipated that continued advances in mobile computing technology will allow more advanced applications in the future. Replacing the existing commercial 802.11 wireless communications technology with the emerging Joint Tactical Radio System (JTRS) and rehosting the Java clients on fielded tactical command and control systems may soon allow a target detection capability at lower echelons. Weather information required as input for these clients is readily available (or will be soon) at echelons down to the battalion level via the fielded IMETS.

REFERENCES

[1] D.K. Wilson, 2000, "Reference Guide for the Acoustic Battlefield Aid (ABFA) Version 2", *ARL-TR-2159,* U.S. Army Research Laboratory, Adelphi, MD.

[2] Sauter, D., Torres, M., 2002, "Mobile Weather Technology for the Army", *Wireless Business & Technology*, February/March 2002, pp. 58-62.

[3] Gouveia, et al, 2000, "TAWS and NOWS: Software Products for Operational Weather Support", *Proceedings of the BACIMO 2000 Conference*, Ft. Collins, CO.

TAWS Ensembles for Use in IWEDA

David Sauter, Richard Shirkey, Barbara Sauter
Army Research Laboratory
White Sands Missile Range, NM 88002 USA

Sean O'Brien
Physical Science Laboratory
New Mexico State University
Las Cruces. NM 88003

ABSTRACT

Providing tactical commanders/users with quantitative detection range information for specific target/sensor pairings and environmental conditions is a desirable capability. However, providing this information for all scenarios (i.e., targets, sensors, locations, etc) in near real time is not realistic given the current processing power of tactical computers and the availability of high resolution environmental data (in both space and time). As a result, for planning purposes and initial screening of potential target acquisition devices (e.g.. IR devices vs. image intensifiers) a simple rule based application can provide valuable guidance. Such an application exists and is fielded via the Army Battle Command System (ABCS) – the Integrated Weather Effects Decision Aid (IWEDA[1]). IWEDA, however, currently has highly simplistic rules that determine whether the particular device will be adversely impacted by the environment. Typically these rules are based on only the meteorological visibility and possibly one other weather parameter (e.g., fog or precipitation). Moreover, there is no target information associated with IWEDA.

In an attempt to provide more realistic weather impact information on target acquisition devices from IWEDA without increasing the access time for the user, an approach is being undertaken that will utilize the quantitative information available from the Target Acquisition Weapons Software (TAWS[2]). TAWS provides detection ranges for a specific sensor given target details and environmental information. By running TAWS for a large number of environmental conditions and targets, parametric ensemble curves can be pre-computed for use in IWEDA. In concert with the simplicity of IWEDA's rules, these curves will provide the user with *average* detection ranges and additional guidance when in the rule-predicted "marginal" (i.e., amber) or "unfavorable" (i.e., red) IWEDA regions. This allows higher fidelity guidance to be provided in the mission planning or operational support stages.

Work is also being completed to allow the IWEDA user to call TAWS functions indirectly from IWEDA for determination of detailed detection ranges using IWEDA user supplied targets and environmental conditions automatically retrieved from a gridded meteorological database.

INTRODUCTION

In an ideal situation, tactical users would be able to instantaneously determine the detection range for a specific target and operating state by a target acquisition device for a number of different times and locations. Having this information at hand would be of great value in both the execution and planning phases of warfare. As stated in the abstract, however, due to the overwhelming amount of environmental data and target/sensor pairs, this is currently not tenable. As a result, the IWEDA application includes some rudimentary rule based information on the performance of various target acquisition devices. This output can be used in the planning of missions to determine which type of device (e.g., visible vs. infrared) may have an advantage over another in acquiring a target. Figure 1 shows IWEDA output for two acquisition devices and the effectiveness over time. The environmental impact(s) for a specific time and system can be retrieved via a left mouse click on a matrix cell. This information is displayed in the figure as well. A matrix cell with "G" is colored green in a color display and represents no adverse impact, a cell with "A" is an amber (marginal) impact and a red ("R") cell indicates an unfavorable impact. In contrast, Figure 2 shows TAWS output in which detection range values in kilometers are displayed over time. Clearly, a product that could combine the efficiency of the rule based IWEDA with the physics based quantitative output of TAWS would be of great value in tactical situations. This desire led to an ongoing effort in which numerous TAWS runs are completed for each of the existing IWEDA sensors against various target types/operating states, line of sight azimuths, geographic location/season and meteorological scenarios. These results will be averaged and provided as functions of parameters such as meteorological visibility, cloud cover and aerosol type. The end result will be a 4th degree polynomial

parametric curve that will provide a normalized average detection range as a function of visibility. For each IWEDA sensor, there may end up being several parametric curves as a function of meteorological conditions (e.g., aerosol type, cloud amount and type, etc). Thus, the IWEDA sensors' performance can be parameterized to include averaged target information as opposed to the non-target dependent version currently fielded. The remainder of this paper discusses some of the initial results of these ensembles and how they may be incorporated into IWEDA.

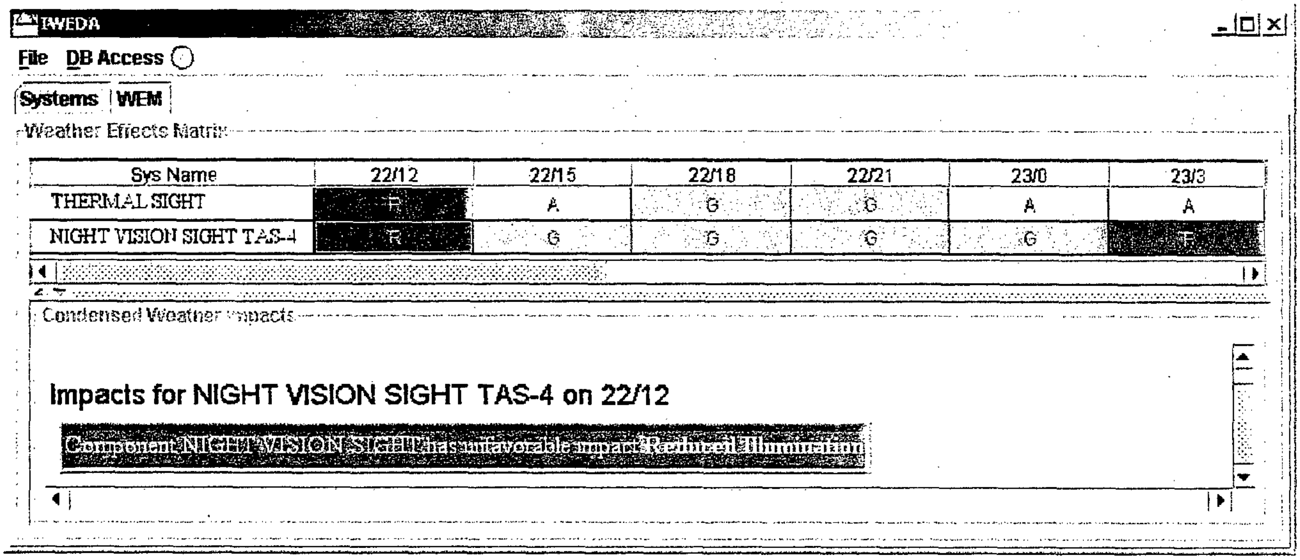


Figure 1 - IWEDA Display for Target Acquisition Devices

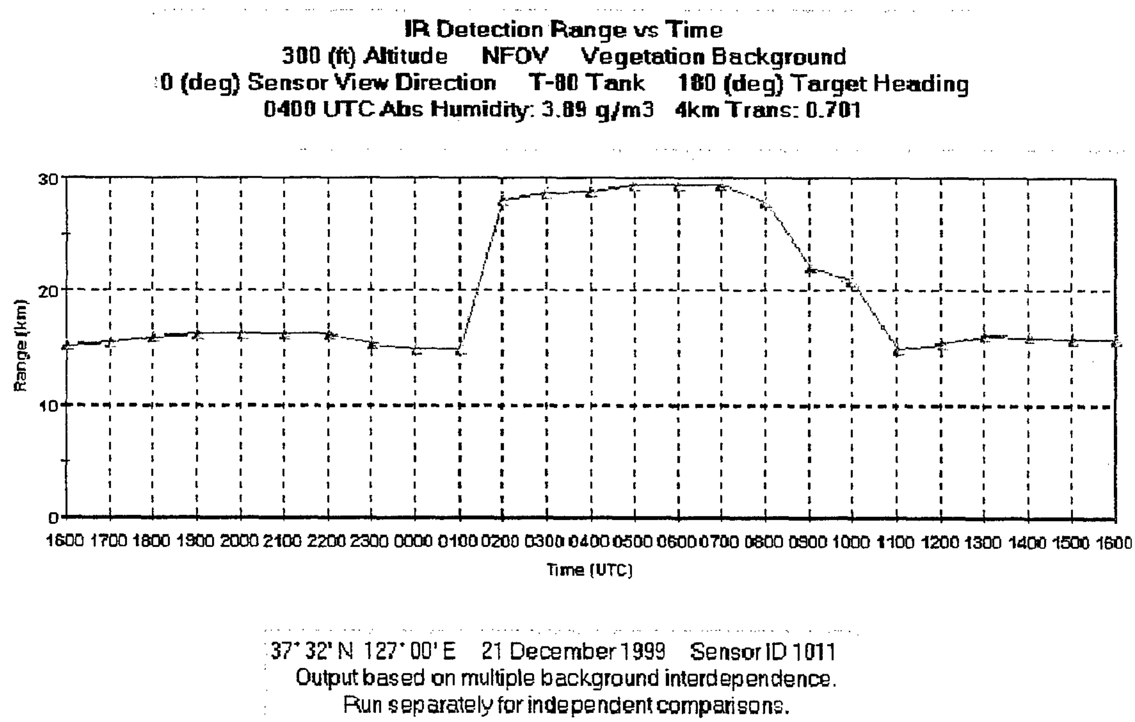


Figure 2 - TAWS Detection Range vs. Time

CURRENT ENSEMBLES

To date, numerous ensembles have been created for 26 narrow field of view (NFOV) IR sensors acquiring a T-80 tank and Armored Personnel Carrier (in both the exercised and off modes) at various times of the day, both with a rural and fog aerosol. The results of these numerous runs have been averaged and fit to several parametric curves as discussed previously. Figure 3 shows the results of some of these ensembles. This figure demonstrates the type of information that will be available for eventual use by IWEDA. Although the curves in the figure are for an ensemble of 26 different IR sensors, the IWEDA ensembles will be computed for the individual sensors corresponding to the current list of available IWEDA sensors. This should produce only a few or several parametric curves for each sensor (as a function of time and/or one or more meteorological parameters, e.g., aerosol type or cloud cover/type). These parametric curves can then be used to determine the normalized average detection range (actual TAWS computed detection range divided by the maximum computed detection range for all conditions). This normalized average detection range will form the underlying basis for the IWEDA favorable, marginal and unfavorable regions as opposed to the current dependence on primarily visibility. Many additional TAWS runs will be required to accommodate the IWEDA sensors and to produce the required ensembles.

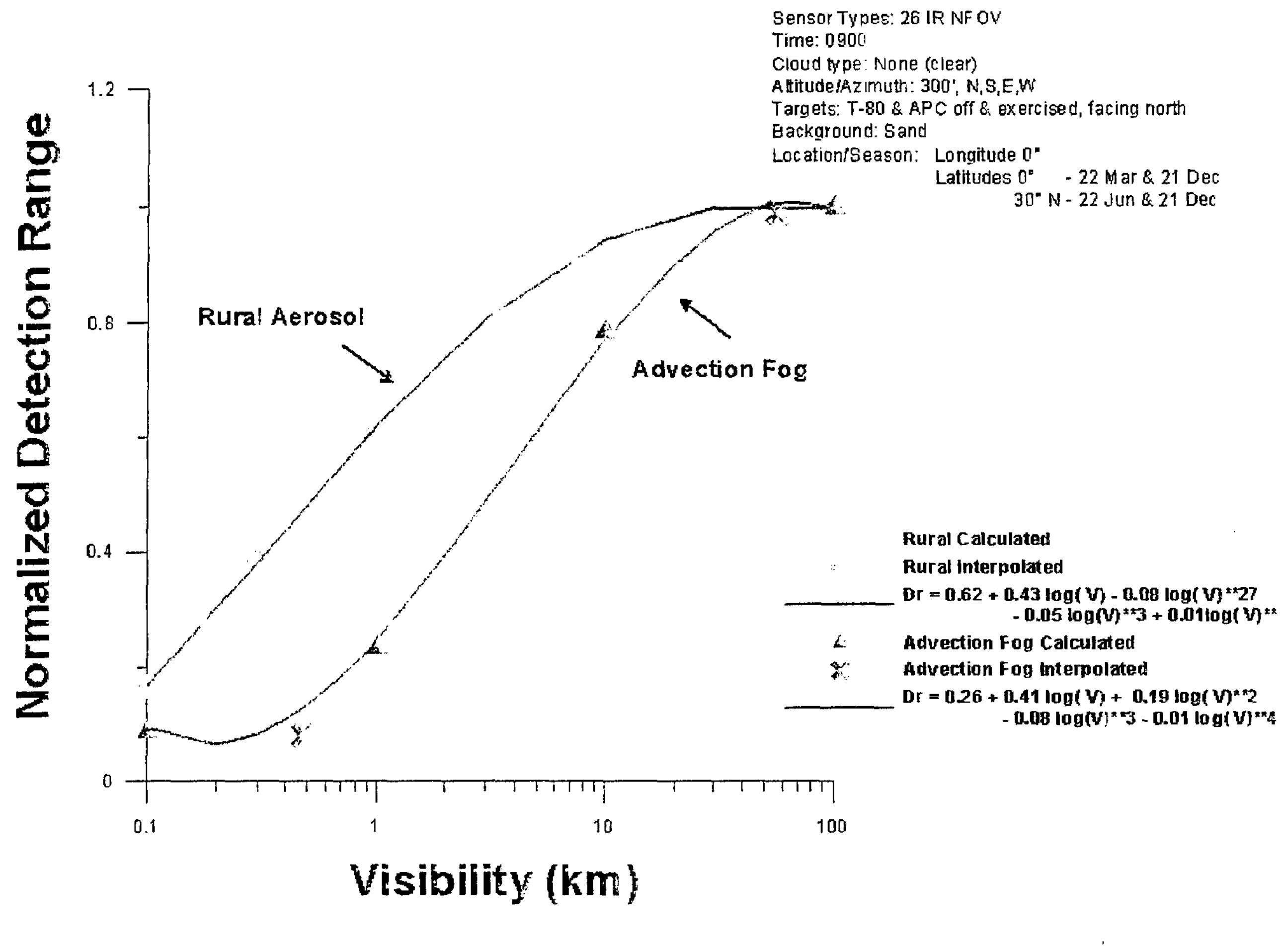


Figure 3 - Rural Aerosol & Advection Fog Parametric Curve Fits

ENSEMBLE INCORPORATION INTO IWEDA

Once the appropriate IR and visible wavelength TAWS runs have been made and the ensembles have been created, the parametric curves for each IWEDA sensor can be plotted and the 4th degree polynomial functions can be determined. There are then two different methods in which IWEDA can incorporate this target dependent normalized average detection range information. These are discussed in more detail in the following sections.

Use of the Polynomial Functions

The normalized average detection range can be computed for each IWEDA sensor and grid point within the IWEDA spatial domain (typically over 2500 points for each forecast period[1]). The polynomial to be used in the computation would be based on the existing meteorological conditions at each grid point. As an example, IWEDA sensor "A" may have a total of 6 polynomial equations to choose from (clear sky, partly cloudy sky and overcast for both rural and fog aerosol). A three-step process is required to determine the IWEDA impact (if any) via use of the polynomial functions. First, retrieve the meteorological parameters for each IWEDA grid point to determine the appropriate polynomial equation to use. Second, compute the normalized average detection range using the polynomial. Lastly, compare this computed normalized average detection value with the threshold value (determined a priori) that has been determined to create a marginal or unfavorable impact. In reality, while the polynomials produce normalized average detection range values, the IWEDA favorable/marginal/unfavorable thresholds will be based on absolute average detection ranges[2]. In this manner. IWEDA will incorporate target information (albeit "average" target signatures) into its output. Until IWEDA sensor ensembles have actually been created, it is unknown how many ensembles will be required for each sensor. The actual number created will be determined somewhat subjectively and will be a function of the magnitude of variability between the parametric curve fits for the various input parameters. For example, if it is determined that cloud amount does NOT significantly affect the average detection ranges for a specific IWEDA sensor, it may be unnecessary to create separate polynomials as a function of cloud cover. Limited results to date for a NFOV IR ensemble average showed little dependence on local time but a significant dependence on whether or not fog was present. This is encouraging, as time of day dependencies would add a level of complexity to the process of determining which polynomial to use. This is due to the fact that the time of day dependence is really a function of the relative azimuth and altitude between the sensor-target viewing direction and the solar position. Thus, for time dependencies, one cannot simply use the time of day for varying geographic locations/seasons but would need to determine the solar azimuth and altitude and then use it accordingly. This is one more computation required for incorporating the polynomial formulation process. Benchmarking will be required to determine the tradeoffs between a large number of ensembles versus increased run time in determining the IWEDA sensor impacts[3]. Obviously the larger the number of ensembles, the better the representation of the meteorological and (any) time of day effects on the average detection ranges. However, having a large number can result in a situation that is overly complex and decreases the IWEDA efficiency and end up defeating the time saving intent of utilizing the ensembles. In the future, the TAWS will be callable from within IWEDA itself, such that the detection range for the IWEDA sensor and user specified target (entered via a graphical user interface that will pop up) may be computed for individual grid points.

Use of the Parametric Curves

A second potential method for utilizing the TAWS ensemble information is to simply use the parametric curves for the individual sensors. In this case, one would also base the IWEDA criteria on the absolute average detection range as previously discussed but would not need to compute the normalized average detection range using the polynomial function. Instead, the graph containing the parametric curve would be consulted a priori to determine the meteorological visibility corresponding to the appropriate normalized average detection range (which again would be determined by mapping an absolute average detection range, e.g., 2.5 km against the maximum average detection range computed for that ensemble) at which values equal to or less would be equated with an unfavorable IWEDA impact (the same would be done to determine

[1] IWEDA provides a map overlay of the spatial variation of the impacts throughout the forecast period (typically a 48 hour prognosis at 3 hour increments). The spatial domain, by default, is a 500x500 km area with horizontal grid spacing of 10 km. This results in a total of 2601 grid points for each forecast period (51x51 grids). The meteorological parameter values at each grid point used by IWEDA are created by a mesoscale model that is run twice a day on the Army's tactical command and control system for weather, the Integrated Meteorological System – IMETS.

[2] Since a *normalized* average detection range of 0.5 for one ensemble may represent an *actual* average detection range of 3 km and in another case 10 km, the *normalized* value does not produce a consistent benchmark between sensors or meteorological scenarios. Thus, a consistent *absolute* average detection range threshold will be determined for the various IWEDA thresholds (e.g., <2.5 km unfavorable, 2.5-5 km marginal and > 5 km favorable) and be applied for all sensors. For IWEDA purposes, the polynomials can either be modified to produce *absolute* average detection ranges or the thresholds can be expressed as a *normalized* value that may vary from sensor to sensor and meteorological scenarios.

[3] Using the previous example of 6 polynomials available for a specific sensor based on cloud amounts and aerosol type, the values for these two meteorological parameters will have to be retrieved from the gridded meteorological database for each of the 2500+ grid points. A decision tree will be coded and utilized to choose the correct polynomial equation to use. This will essentially be a set of "if-then" programming languages statements (e.g., "if cloud amount = 0 and aerosol = fog then use polynomial X"). Obviously the more meteorological parameters that are used as the basis for individual polynomial equations, the lengthier the decision tree (and the longer the sorting and computation process, particularly when one considers the cumulative effect over 2500 grid points).

the corresponding visibility or visibility range for the marginal impact). Using Figure 3 as an example, assume for the moment that these two curves represent the response of IWEDA sensor "A" to visibility for a fog aerosol and a rural aerosol. Also assume that the maximum average detection range computed for all of the ensembles was 12.5 km. Then the unfavorable impact for this sensor (from footnote 1, an absolute average detection range of 2.5 km) would be represented by an average normalized detection value of 0.2 (2.5 km/12.5 km) or less. Reading the corresponding visibility value for this normalized average detection range for the advection fog case yields a value of approximately 1.0 km and just over 0.1 km for the rural aerosol case. This provides a simpler and faster determination of the IWEDA impact level than using the polynomial equations. as it is neither necessary to determine which polynomial to use nor compute the detection range using the polynomial. In this case, the actual average detection range of sensor "A" is not available for display to the IWEDA user as it would be if the polynomial equation was used. However, as mentioned previously, it is anticipated that the IWEDA user will soon have the capability of computing the actual detection range for specific grid points on demand, thus this shortcoming is deemed non-critical in light of the increased efficiency (benchmarking will be required to determine whether the computing time difference is significant).

To date, only NFOV IR sensor ensembles have been produced. IR sensors, in general, should be less sensitive to the relative angle between the sensor-target viewing direction and the solar position. For visible sensors, however, it is anticipated that the angular dependence (i.e., time of day) will be more significant, particularly for cases when the sun is unobscured. This is due to the fact that the sun will cause target shadows in the visible wavelength. Reflected light off target facets may also contribute to the angular dependence. Careful thought will need to be given as how to best represent visible sensor ensembles without overly complicating the situation and creating a large number of ensembles.

SUMMARY

Work has been initiated to incorporate target effects into the rule based IWEDA application via an ensemble approach. This effort has the potential to provide more realistic guidance to the IWEDA user regarding the environmental impacts on the current inventory of IWEDA sensor devices. A potential drawback is the possibility of making the underlying IWEDA application more complex if too many parametric ensembles are added. On the other hand, if too many parameters are included in the individual ensembles (resulting in relatively few ensembles), a danger exists of simply smoothing out the target effects resulting in output that has added little if any value to the situation. The ultimate solution is to utilize TAWS to compute the actual detection ranges for each of the multitude of IWEDA grid points (using the unique meteorological values) for a user specified target for each sensor. Due to tactical computing limitations, however, this capability is some years off. In the interim, allowing the user to input a specific target for an individual IWEDA grid point via a simple GUI to determine the detection range will provide additional information. Coupled with the initial guidance on the effectiveness of the sensors from the parametric curves, this will hopefully provide a satisfactory solution.

REFERENCES

[1] Sauter, D., et al, 2001, "Leveraging Command and Control Technology to Provide Advanced Environmental Effects Decision Aids on the Battlefield", *Proceedings of the 5th International Command and Control Research and Technology Symposium*, Canberra, ACT, Australia.

[2] Gouveia, et al, 2000, "TAWS and NOWS: Software Products for Operational Weather Support", *Proceedings of the BACIMO 2000 Conference,* Ft. Collins, CO

Modeling and Validation of the Effects of a Metal Ground Plane on the RCS of an Asymmetric Trihedron

William A. Spurgeon
U.S. Army Research Laboratory
ATTN: AMSRL-WM-MB
Aberdeen Proving Ground, MD 21005-5069

Christopher Kenyon
U.S. Army Research Laboratory
ATTN: AMSRL-SE-RU
Adelphi, MD 20783-1145

Abstract

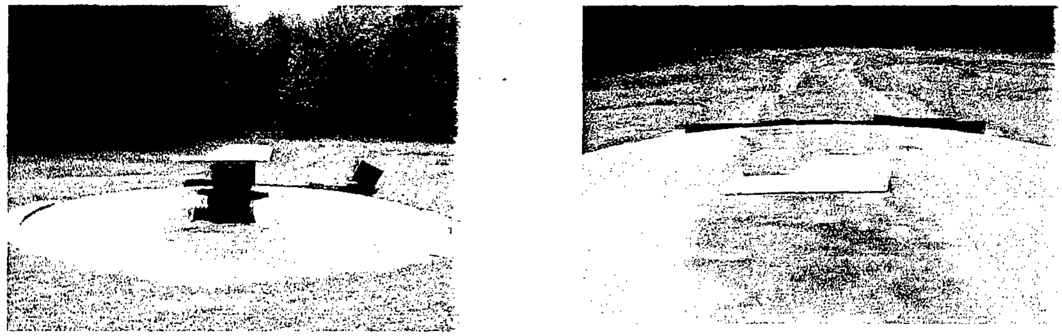
The T5M3 trihedron is a small and physically simple radar target that was used for modeling and experimental studies of the effects of a metal ground plane on its radar cross section (RCS). This trihedron is asymmetric and has one vertical face that is slanted backward 15 degrees from the vertical. Despite its physical simplicity as a radar target, it requires an exact solver code to accurately predict its RCS. In outdoor experimental tests, it was found to exhibit a significantly different RCS depending upon whether it is placed directly on the turntable or is isolated from the turntable. This paper presents the results of model RCS calculations via Xpatch, an approximate, physical optics-based, shooting-and-bouncing rays code, and a Method of Moments code (the fast Illinois solver code or FISC), as well as the experimental results. These calculations show the mechanisms by which the RCS of the trihedron is impacted by the metal turntable.

Introduction

This paper presents an interim report on our investigations of the interaction of a structurally simple, but electrically complex, radar target with a ground plane. This target is shown on the U.S. Army Research Laboratory (ARL) radar range turntable in Figure 1. It is an asymmetric trihedron, about 54 inches long by 24 inches high and 8 inches high, with the shorter upright side canted backward 15 degrees from vertical. The difficulties in modeling this target have been described in detail by Carter [1]. Experiments were conducted to obtain valid data. Modeling results were obtained from Xpatch* [2], an approximate, high frequency shooting-and-bouncing rays code, and from an exact integral equation (Method of Moments)[3] solver, the Fast Illinois Solver Code (FISC).*

---

* SAIC-DeMaco, 100 Trade Center Drive, Suite 303, Champaign, IL, 61820.

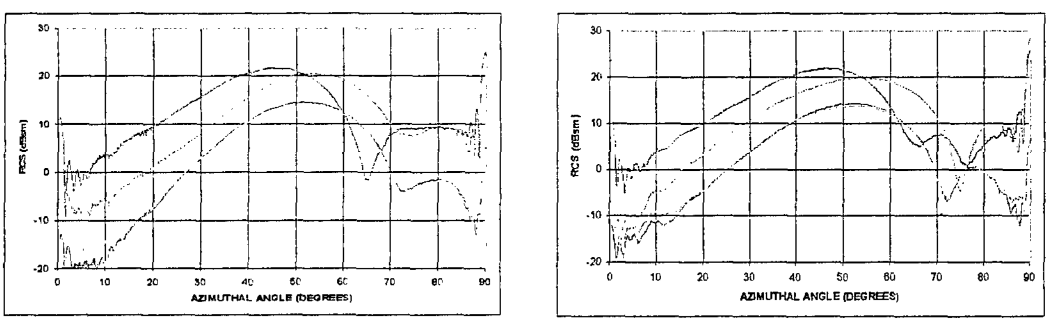


(a) (b)

Figure 1. The trihedron shown on the turntable (a) and isolated from the turntable on a 38-inch-high wooden table surrounded by 4-inch pyramidal foam radar absorber (b). The view is looking away from the radar.

Experiments and Results

The trihedron was fabricated in the ARL machine shop from tool and fixture aluminum with a 25-microinch finish and a flatness of 0.015 inches across a 4- by 8-foot sheet. The sides were bolted together to ensure good electrical contact. All tests were conducted at the ARL outdoor radar range on a clear day in August 2001. The radar was set for a center frequency of 9 GHz and had a 1.6-GHz bandwidth. The radar was on a tower about 150 feet from the target, so the strongest scatterers on the target were in the far field. A 10-degree depression angle was used for all tests. The RCS of the empty turntable was generally well below that of the trihedron. The trihedron was tested both directly on the turntable and on a 38-inch-high wooden table covered with 4-inch pyramidal radar absorber (Figure 1). Experimental results for these two arrangements are shown in Figure 2.



(a) (b)

Figure 2. The measured 9-GHz RCS of the trihedron isolated on the turntable (a) and on the wood table (b). Legend: blue = VV; red = VH; orange = HV; green = HH.

The major differences between the two sets of measurements areas are as follows.

1. The minimum in the VV channel (blue curve) for the isolated trihedron at an azimuthal angle of about 65 degrees becomes a double minimum for the trihedron on the turntable.

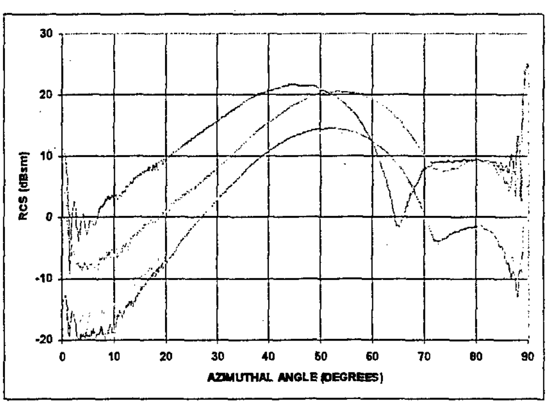
2. The minimum in the HH channel (green curve) at an azimuthal angle of about 74 degrees is about 13 dBsm deeper.

3. The minima in the cross-polarized channel RCS (VH or HV channels, red or orange curves) are a few decibels relative to square meter deeper.

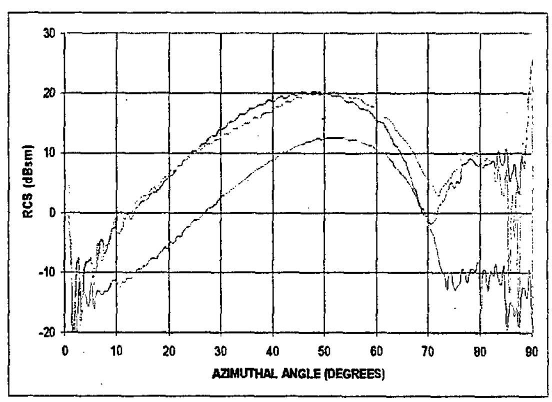
The VH and HV channel RCS values should be identical and the small extent to which they are different indicates that this is good data. Additional experimental details will be presented in a forthcoming ARL Technical Report [4].

Xpatch Calculations

Xpatch is an approximate shooting-and-bouncing rays code that was developed for calculating the RCS of large targets at high frequencies. As an approximate code, it might be expected to miss some fine details in an RCS calculation, but it should derive the main features reasonably well and be able to predict how the RCS is changed as depression angles or frequencies are changed, or if a target is altered by adding a ground plane. The 9-GHz data and calculated Xpatch results for the isolated trihedron are shown in Figure 3.



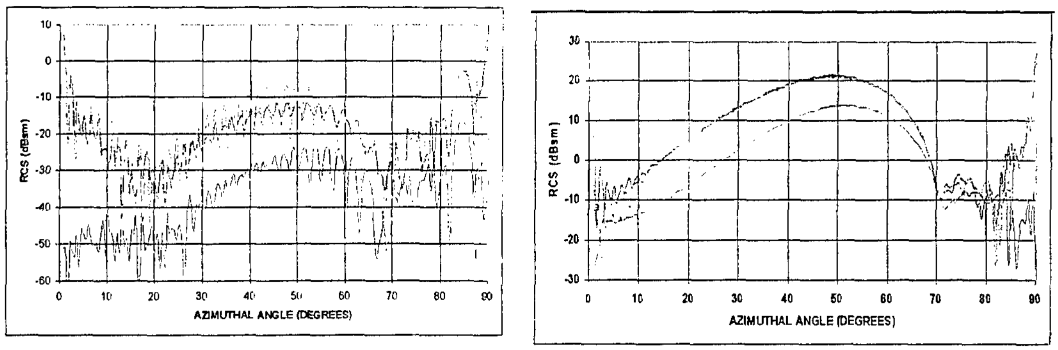
(a)



(b)

Figure 3. Xpatch calculation results (a) and experimental data (b) for the isolated trihedron. Legend: blue = VV; red = VH; orange = HV; green = HH.

The Xpatch results for the HH and VH channel RCS are in reasonable agreement with experiment, but the VV channel RCS differs significantly from the data (Xpatch calculates the VH and HV channel RCS, then averages the two [5], thus obtaining reciprocity). Additional Xpatch calculations indicate that the first bounce contribution to the RCS is small except at 0 and 90 degrees, and the large peak in all channels arises on the second bounce between the two vertical sides. These results are shown in Figure 4.



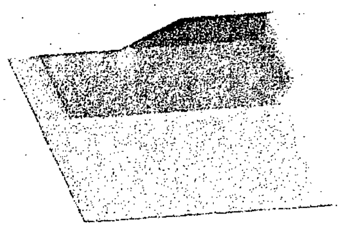
(a) (b)

Figure 4. The first bounce (a) and two bounce (b) contributions to the RCS. Legend: blue = VV; red = VH; orange = HV; green = HH.

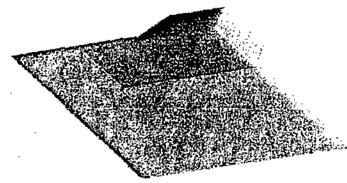
The major differences between the data and the Xpatch results, particularly the difference between the VV and HH channel RCS, arise on the third bounce. The main reason for this is that the third bounce effects are very complicated at this depression angle, making tracking phase and amplitude of the reflected rays difficult. The geometrical arrangement is also such that a traveling wave that could not be modeled by Xpatch may be a part of the problem [6].

Four experimental arrangements were modeled using Xpatch. These are shown in Figure 5. The first is for the trihedron on a 62-inch square plate. If a ray is to hit one of the vertical sides of the trihedron and miss the base, altering the RCS by a direct hit, it would have to hit the turntable ground plane within 47.5 inches of the vertical seam on the trihedron. The 62-inch plate is wide enough for this with a small margin on the sides. The second ground plane arrangement is an 80-inch-wide ground plane. Any ray that can pass through the bounding box for the trihedron must hit within 71.5 inches of the seam. Such rays could alter the RCS by changing the phase of the radiation that is reflected on the third bounce. The third is a 20-foot-diameter circular plate as shown in Figure 5. In a ray tracing explanation of the effects of the ground plane on the RCS of the trihedron, this should be no different from

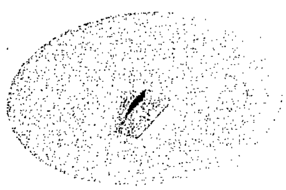
the 80-inch plate result. If one considers the RCS as arising from induced currents that reradiate, however, the RCS results for the larger ground plane would be expected to differ from the 80-inch plate result, but the magnitude of the difference would have to be calculated. Finally, the trihedron on the wooden table surrounded by a perfect absorber was modeled as shown.



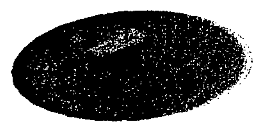
(a)



(b)



(c)

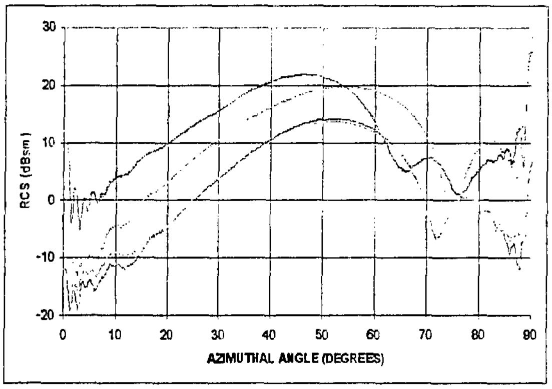


(d)

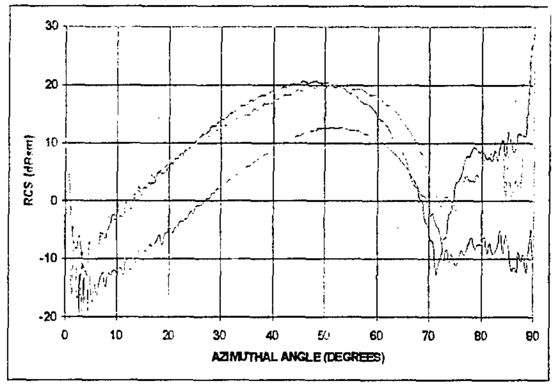
Figure 5. The four experimental variations modeled via Xpatch. Legend: (a)62-inch plate; (b) 80-inch plate; (c)20-foot-diameter plate; (d) isolated on wood table.

The experimental data and the Xpatch results for the first three models are shown in Figures 6. For the 62-inch square ground plane, the Xpatch results show a much deeper VV minimum at about 70 degrees azimuth, but not a double minimum. The HH minimum is also deeper, but not as deep as for the data, and the cross-polarized channel results are a little off near the minimum. For the 80-inch square ground plane, the minimum in the VV and HH channels are broader, and there is no double minimum in the VV channel RCS. The cross-polarized channel RCS is in reasonable agreement with the data except for azimuthal angles larger than about 80 degrees. For the 20-foot-diameter ground plane, the VV minimum is at about 68 degrees and is still a broad single minimum rather than a double minimum, the HH channel results are closer to the experimental values, and the cross-polarized channel RCS is in fairly good agreement with the data except for azimuthal angles greater than

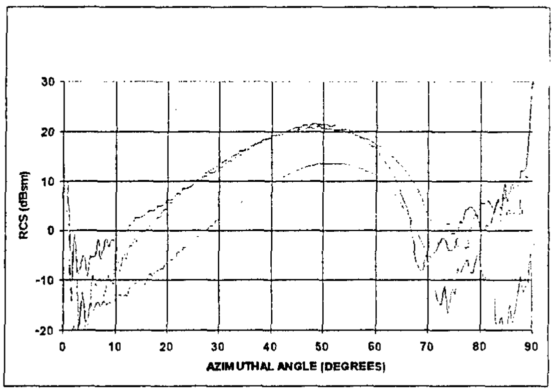
about 78 degrees. In all cases, the HH and cross-polarized channel RCS are in reasonable agreement with experiment for azimuthal angles below 50 degrees, while the VV channel experimental data differs from the calculation by 10 dBsm.



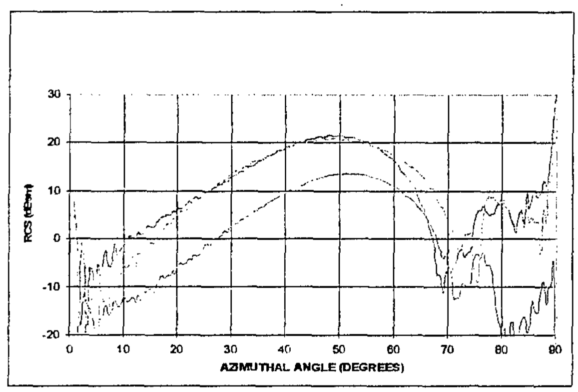
(a) Experimental Data



(b) 62-inch Ground Plane



(c) 80-inch Ground Plane



(d) 20-foot-Diameter Ground Plane

Figure 6. The experimental data (a), Xpatch results for a 62-inch square ground plane (b), Xpatch results for an 80-inch square ground plane (c), and Xpatch results for a 20-foot-diameter ground plane (d). Legend: blue = VV; red = VH; orange = HV; green = HH.

Xpatch should obtain the same result for the isolated trihedron and for the trihedron on the wooden table if this arrangement does isolate the trihedron from the turntable. Figure 7 shows the calculated VV channel results for the isolated trihedron and for the trihedron on the wood table. The difference is small as expected. Similar results were obtained in the other channels.

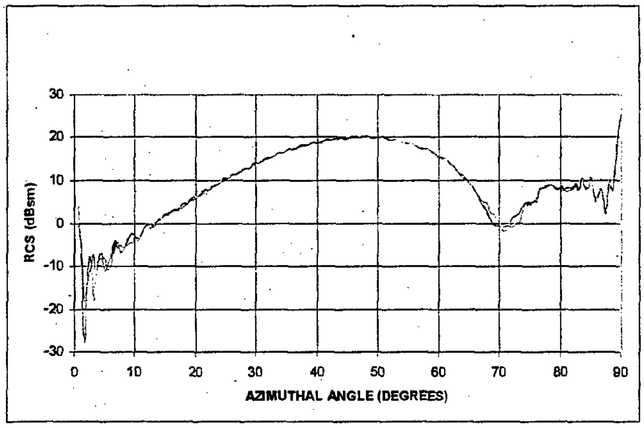


Figure 7. VV channel Xpatch results for the completely isolated trihedron (blue curve) and for the trihedron on the wood table (red curve).

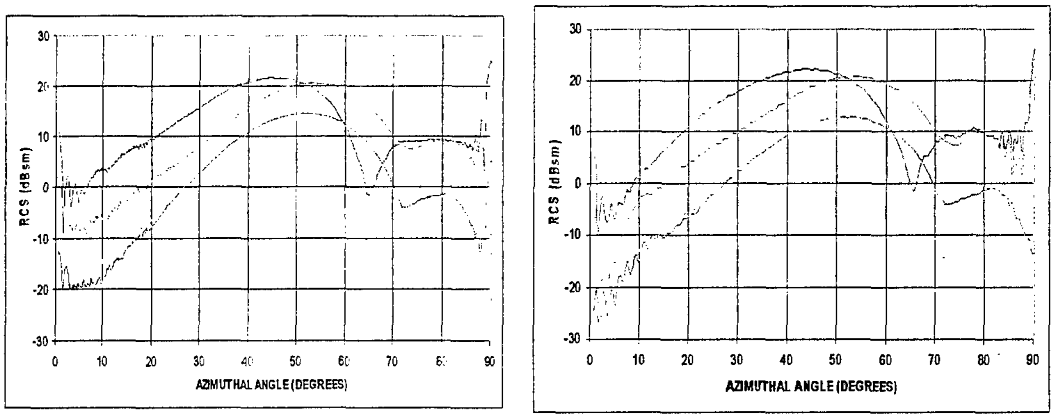
FISC Results

Modeling calculations were also performed using FISC, a nominally exact code that should result in the correct answer. However, it can be set to give results of low, medium, or high accuracy as a modeler's option; it was set to medium accuracy for the calculations reported here. One other option is the use of the electric field integral equation (EFIE) or the magnetic field integral equation (MFIE) or some combination of the two (combined field integral equation or CFIE) to solve the problem. The most frequent choice, and the one appropriate for a closed electrical body such as the trihedron, is to use a CFIE composed of one-half the EFIE and one-half the MFIE; this selection was used in the calculations.

FISC uses the same facet model as Xpatch, but refines the facets so that facet sides are less than 1/10 wavelength long. It also has an option to include an infinite metal ground plane under a target. A problem will run about 2 or 3 times slower that the corresponding problem with an isolated target. The Xpatch results suggest that the entire turntable contributes to the RCS of the trihedron on the turntable, and the infinite ground plane option was selected as being a better a better model than the smaller ground planes previously described.

The experimental data and the FISC results for the isolated trihedron are shown in Figure 8. Unlike the Xpatch model results, The FISC results agree well with experiment in the azimuthal angle range from 50 to 85 degrees, where the Xpatch results did not agree well at all; in particular,

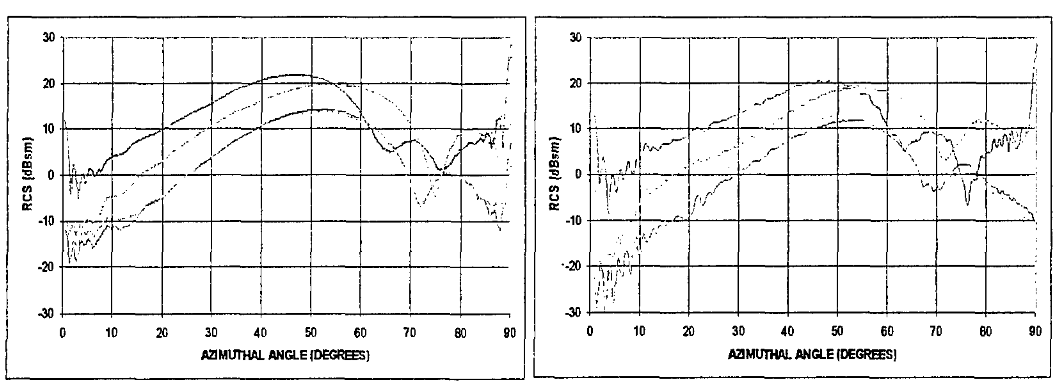
the difference in the VV and HH channel RCS is predicted. The major differences between the FISC results and the data are in the cross-polarized channel between about 87 and 90 degrees and the HH and cross-polarized channels (and, to a lesser extent, the VV channel) from 0 to 10 degrees. The VH and HV results are generally fairly close; they should be identical.



(a) (b)

Figure 8. The experimental data (a) and the FISC calculation results (b) for the isolated trihedron. Legend: blue = VV; red = VH; orange = HV; green = HH.

The experimental data and FISC results for the trihedron on the turntable are shown in Figure 9. As for the isolated trihedron, the FISC results are in good agreement with the data for all azimuthal angles except for the 0 to 10 and 87 to 90 degrees ranges. The double minimum in the VV channel and the deepening of the minimum in the HH channel are reasonably well predicted. The difference between the VH and HV results is a little greater than for the isolated trihedron.



(a) (b)

Figure 9. The experimental data (a) and FISC calculation results (b) for the trihedron on the turntable. Legend: blue = VV; red = VH; orange = HV; green = HH.

Conclusions

Xpatch predicts a definite interaction between the ground plane and the trihedron, but misses fine, but important, details, probably because of the approximations it uses. It does predict that the trihedron on the wood table is fairly well isolated from the ground plane. The calculations are also fast, requiring 10 to 120 minutes of computation time depending on the size of the problem. FISC with medium accuracy does a much better job of predicting the results and gets many details, such as the double minimum in the VV channel and the differences in the VV and HH channels. Still, it does not get all details correct; the results for azimuthal angles between 0 and 10 degrees and 87 and 90 degrees differ from the data by as much as 10 dBsm. The high accuracy option in FISC should give a more accurate prediction, but will take even longer than the 1000 plus hour computation time with the medium accuracy option.

Future Plans

Additional measurements of the trihedron both isolated and on the ground plane at frequencies of 10, 35, and 94 GHz and depression angles of 10, 15, 20, and 30 degrees are planned. Additional modeling studies using Xpatch, FISC, or a finite difference time domain code will also be pursued. Other codes that should provide accurate results for this problem are under development and will be tried when they become available. It will be particularly interesting to know how well Xpatch can model these results at other depression angles and higher frequencies.

Acknowledgments

The authors would like to thank ARL colleagues R. Tan and S. Stratton for their help with the experiments, D. Hopkins for help with the facet files, and P. Dehmer for reviewing the manuscript. Special thanks are due S. Carter from the National Ground Intelligence Center for sharing this problem with us.

References

1. S. L. Carter, "T5M3 Trihedron RCS Anomalies," NGIC Report 1141-03231-00, National Ground Intelligence Center, Charlottesville, VA, December 1999.

2. D. J. Andersh, M. Hazlett, S. W. Lee, D. D. Reeves, D. P. Sullivan, and Y. Chu, "XPATCH: A High Frequency electromagnetic Scattering Prediction Code and Environment for Complex Three-

Dimensional Objects," *IEEE Antennas Propagation Mag.*, vol. 36, pp. 65-69, February 1994.

3. E. F. Knott, J. Shaffer, and M. T. Tuley, *Radar Cross Section*, 2nd ed., Artech House, Boston, MA, 1993, pp. 124-139.

4. W. A. Spurgeon, S. R. Stratton, and R. J. Tan, "Outdoor Radar Range Radar Cross-Section Measurements on the T5M3 Trihedron," U.S. Army Research Laboratory, Aberdeen Proving Ground, MD, to be published.

5. S. Kasanovitch, private communication, SAIC-DeMaco, Champaign, IL, 26 January 2002.

6. Knott et al., op. cit., pp. 228, 242.

Advances in Modeling Visual Search and Target Discrimination Performance

G. Witus
Turing Associates, Inc.
1392 Honey Run Drive, Ann Arbor, MI 48103

R.E. Karlsen, G.R. Gerhart, and D.J. Gorsich
U. S. Army TARDEC
Warren, MI 48397

Abstract

This paper reports on advances in mathematical models of observer-ensemble performance in narrow-field-of-view visual search and target discrimination for ground vehicles in natural terrain. Three developments are presented. We show that the distribution of search time follows a lognormal distribution. We show that search outcome is the result of a race between scene parsing and target detection. Scene parsing guides and focuses search, but also leads to quitting without detection. Quitting is not simply the consequence of having exhausted the supply of suspect locations. We present a refined target signature metric and computational method that emulates human perceptual organization of the target into component regions. This is significant not only because it improves the ease-of-use and reduces subjective user input, but also because it is potentially applicable to thermal images. The refined metric provides reasonably accurate prediction of probability of detection for cued detection and uncued search experiments.

1 Introduction

This paper reports on advances in mathematical models of observer-ensemble performance in narrow-field-of-view visual search and target discrimination for ground vehicles in natural terrain. These results build on and extend visual target discrimination performance modeling and perception experiments previously reported [1].

Three developments are presented. Section 2 shows that the distribution of search time follows a lognormal distribution, in contrast to the commonly-used negative exponential model. Section 3 shows that search dynamics have the characteristics of a race process between target detection and quitting based on parsing the scene into recognized non-target objects or regions. Section 4 presents a refined signature metric and computational method that emulates pre-attentive perceptual organization of the target. Section 5 discusses the implications and limitations of these advances.

2 Distribution of Search Time

The basic theory is that search involves two complementary activities: (1) recognizing background objects, and (2) examining suspect target locations. Recognizing background regions guides and focuses search by eliminating areas from further consideration, indicating promising locations to search, and by guiding expectations of apparent target size. When the scene is fully parsed into recognized background objects (to the observer's satisfaction), search is terminated by quitting without reporting a detection.

Search begins with a scan of the image to begin parsing and detect highly obvious targets. As the scene is parsed, background objects and regions are recognized and rejected (e.g., the sky, distant hillsides, close and empty foreground fields). Suspicious locations, when noticed, are examined. Eye movement studies have found evidence of two states in visual search [2].

The observer's thresholds to examine a suspect location and to decide to designate a target begin high and are lowered over time. This helps ensure that low-confidence suspect locations do not delay or pre-empt inspection and detection at higher confidence locations.

This dynamic implies that the hazard rate, i.e., the rate of search termination as a function of time given that search is still in process, has a characteristic shape:

- It begins at zero.
- It increases as scene parsing guides and focuses search.
- It tails off to zero as the supply of suspect locations is exhausted.

The hazard function for the lognormal distribution has the shape characteristics required by the search theory. The hazard function of the negative exponential distribution is a constant. Neither do generalizations of the negative exponential distribution, e.g., the Weibull distribution and Gamma distributions, have the required hazard function shape. The lognormal model has been used as a model of human information processing in other contexts [3].

Table 1 tabulates several measures of the goodness of fit between the empirical search time distribution and model distributions, for the distribution of search times for 1151 images. Each empirical search time distribution consists of 23 points, i.e. the search times for 23 observers. Four model distributions are compared: lognormal, negative exponential, Weibull, and the sum of two different negative exponential (i.e., a negative exponential delay followed by a negative exponential search time). The measures of fit are the root-mean-square (RMS) error in the cumulative distribution, the proportion of variance (PV) in the empirical distribution explained by the estimate, the maximum Kolmogorov-Smirnov (K-S) statistic over all 1151 images, the average K-S statistic, and the linear regression slope and intercept. The K-S statistic is a measure of the difference between continuous distributions (whereas the chi-squared distribution is used for discrete distributions). The K-S statistic is the maximum of the absolute value of the difference between the two cumulative distribution functions. The lognormal model is a better match than the other model by all measures of performance.

	Lognormal	Neg. Exponential	Sum 2 Neg. Exp.	Weibull
RMS Error	0.068	0.085	0.079	0.084
PV Explained	0.940	0.888	0.917	0.908
Maximum K-S	0.276	0.393	0.357	0.389
Mean K-S	0.127	0.168	0.145	0.152
Slope	1.000	0.823	0.972	0.711
Intercept	0.007	0.125	0.012	0.050

Table 1: Comparison of Search Time Distribution Models

The results of applying the W test of normality of a distribution [4] at different significance levels to the log search time for each of the 1151 images are summarized in Table 2. The significance level is the probability of rejecting the hypothesis of normality when the data are drawn from a normal distribution. The results indicate that the hypothesis of normality is rejected for seven percent of the images, and is questionable for another eight percent.

Percent Rejection for Samples of a Normal Distribution	Percent Rejection for Image Search Time Distributions	Difference
50%	56.9%	6.9%
40%	51.6%	11.6%
30%	43.3%	13.3%
20%	34.5%	14.5%
10%	25.0%	15.0%
5%	17.5%	12.5%
1%	8.3%	7.3%

Table 2: Results of the W Test for Normality of Log Search Time Distributions

The lognormal distribution has two parameters: the mean log time and the standard deviation of log time. In contrast, the negative exponential distribution has only one parameter: the rate which is equal to one over the mean time. The proposed application of the lognormal model uses one constant value of the standard deviation of log search time for all images, set equal to the mean over all images standard deviation of log search time (0.83). Thus the proposed lognormal model requires one parameter for each image (as does the negative exponential model), plus one additional constant for all images.

3 Race Between Quitting and Target Detection

The proposed model asserts that two processes are active during search: recognition and rejection of background objects/regions (exclusive processing) and location and examination of suspect regions (inclusive processing). Individuals employ a mix of these two strategies. Observers relying more on inclusive processing will tend to quit only after exhausting the supply of suspect locations, and will tend to have longer quitting times than detection times. Observers relying more on exclusive processing will tend to take less time to come to the decision to quit than to detect difficult targets. Obvious targets may be detected during the scene parsing. In general, observers will employ a mix of these strategies. They may shift between strategies during search of a single scene, and may change emphasis from scene to scene.

Data from the search experiment supports the proposed model. Figure 1 shows that some subjects tend to have longer quitting times than detection times, while other subjects tend to have shorter quitting times than detection times. (The figure also shows that the search time for low-confidence detections is consistently longer than the search time for high-confidence detections.)

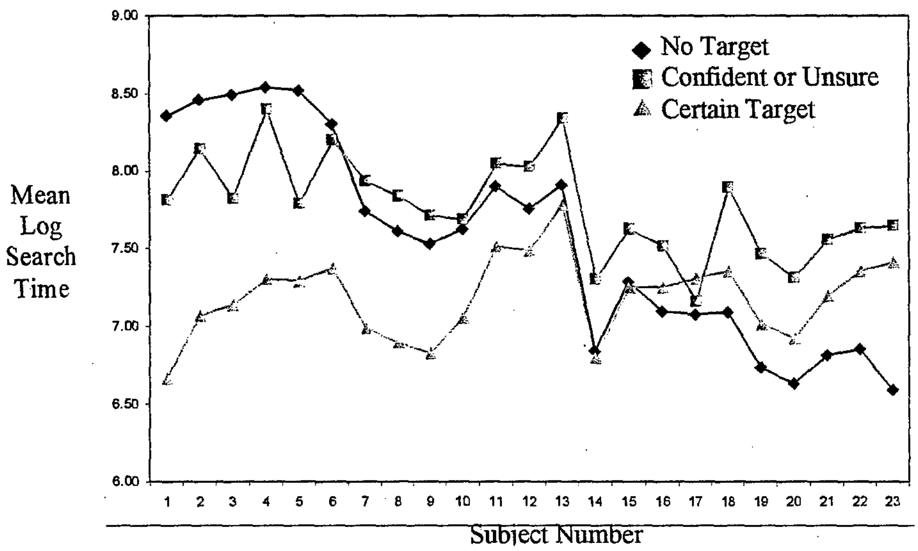


Fig. 1: Log Mean Search Time by Response Rating by Subject

Wolfe's Guided Search model [5] holds that pre-attentive processing builds up a queue of suspect locations whose activation level (based on similarity to the targets) is above the observers threshold, which are then inspected in order of decreasing activation. In the Guided Search model, quitting occurs when the list of suspect locations is exhausted. This implies that the time-to-detect will tend to be shorter than the time-to-quit. The experimental data contradicts the exhaustive search model.

The proposed model implies that the quitting and detection outcomes are in a race: the observer reports a detection or quits depending on whether he detects a target before parsing the scene into recognized background regions and objects. But it is not an independent race. Rejecting background regions makes progress towards quitting, but also guides and focuses search. Noticing suspect locations to examine contributes to quitting when they are rejected, but also contributes to the target detection outcome.

In a race process, the probability of detection for a target-in-background image can be expressed as a function of the search time for the target-in-background image and for the corresponding background-only image. Search time for a target-in-background image can be expressed as a function of the probability of detection and the search time for the corresponding background only image. The probability of detection is the probability that a random variable having a lognormal distribution with mean and standard deviation of log search time for the target-in-background image, μ_{TB} and σ_{TB}, is less than a random variable having a lognormal distribution with mean and standard deviation of log search time for the background-only image, μ_B and σ_B.

$$P_d = 2 - 1 / N\left[\frac{\mu_B - \mu_{TB}}{\alpha(\sigma_B^2 + \sigma_{TB}^2)^{1/2}}\right] \tag{1}$$

where P_d is the probability of detection for the target-in-background image, N and N^{-1} are the standard normal distribution and its inverse. The free parameter α represents the effects of correlation between the processes leading to quitting and detection. Its value, estimated from the empirical data, is 0.6. A value of unity would have indicated complete independence, whereas a value approaching zero would have indicated near perfect correlation.

The equation for P_d as a function of search time parameters can be inverted to compute the mean log search time for the target-in-background image. The standard deviation of log search time for the target-in-background image is replaced with a constant, β, equal to the average value of σ_{TB} over the set of target-in-background images.

$$\mu_{TB} = \mu_B - \alpha\,(\sigma_B^2 + \beta^2)^{1/2}\,N^{-1}\left[\frac{1}{2 - P_d}\right] \tag{2}$$

Figure 2 shows a scatter plot of the experimental estimate of the probability of detection plotted against the probability of detection predicted by the search times and the race model. The 1151 images were organized into 88 groups. Each group represents all of the variations of the 44 base images. However, variations with the unaltered, baseline targets, and the modified, low-contrast targets, were not mixed. The results show that the race model produces an accurate estimate of probability of detection.

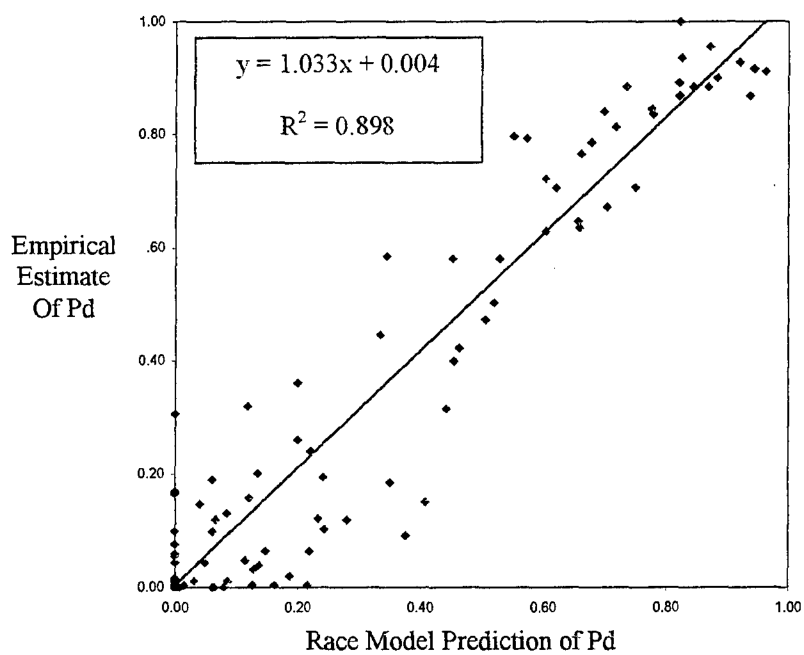


Fig. 2: Race Model Predictions of Pd vs. Empirical Estimate of Pd

Figure 3 shows the scatter plot of the experimental versus predicted mean log search time. The model predictions are highly correlated with the observed data, but the model does not provide as accurate an estimate of search time. The model prediction accounts for only 68.5% of variance in the observed mean log search times.

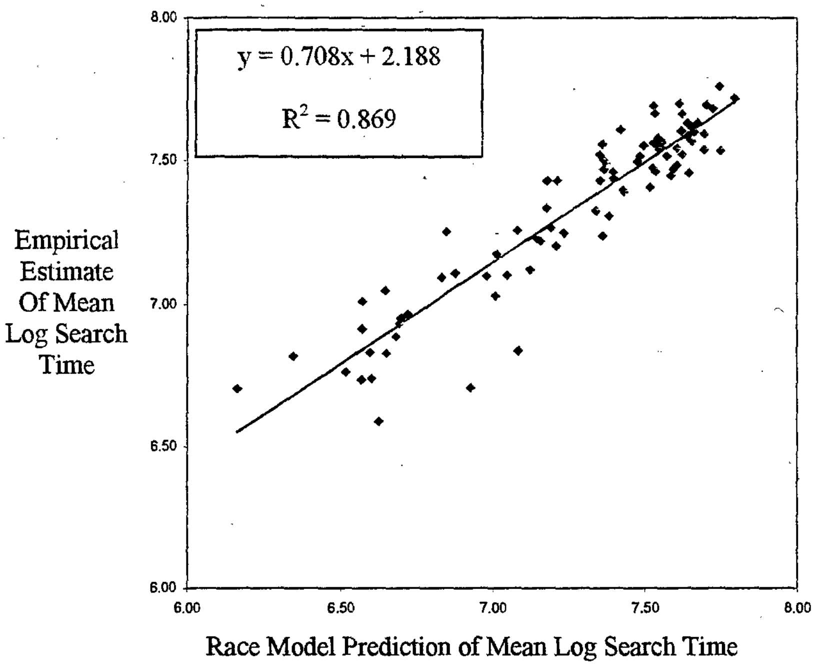


Fig. 3: Race Model Predictions vs Empirical Estimate of Mean Log Search Time

Figure 4 shows a scatter plot of the experimental mean log search time versus the empirical estimate of the probability of detection. The data show that the simple linear regression model provides as good an estimate of mean log search time as does the race model.

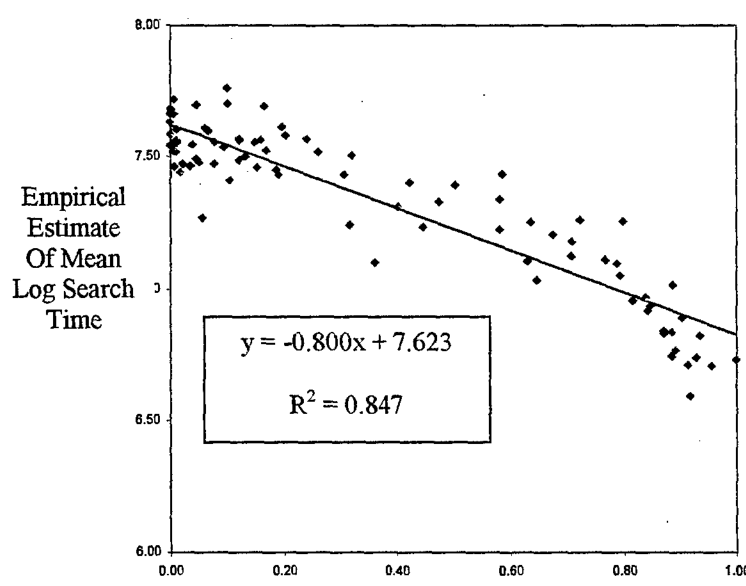


Fig. 4: Empirical Estimate Relationship between Pd and Mean Log Search Time

4 Target Signature Metric

The proposed theory of figure-ground segregation in target detection holds that pre-attentive vision organizes the target into either one region (a silhouette) or two regions (a light or illuminated region and a dark or shadowed region). Visual perception organizes the target into a dominant region and its complement. The dominant region is the region of the target with the greatest distinctiveness. The complement is the difference between the entire target and the dominant region. The perceptual organization does not depend on the combined distinctiveness of the two regions. Attentive visual processing imposes overall target bounding, and categorizes the object based on the joint contribution of the primary region and its complement.

The distinctiveness of a region is a function of the local contrast across the region boundary, the local clutter in the vicinity of the boundary (the variation within the region), and the size of the region. Local contrast and clutter are a function of the region size and shape. Region distinctiveness does not incorporate shape information value. Shape information contributes to attentive discrimination, but not pre-attentive segmentation.

The value of the signature for attentive object discrimination is a function of the combined region distinctiveness, the combined shape information magnitude, and the combined region area. It is not the sum of the value from the two regions, but the value of the combination of the two complementary regions.

The previously-developed target signature metric [1] required the user to define the canonical front, side and top regions of the vehicle. The rationale was that these regions defined the 3D structure of the vehicle, and they defined the 3D appearance of the vehicle since they had significantly different surface normal vectors with respect the source of illumination.

In practice, there were difficulties with this approach. It required a complex set of rules to deal with rounded turrets, shadows, and flat facets that were in the cardinal planes. It increased the user burden, and increased the amount of user subjectivity. Furthermore, the region organization was appropriate only for visual images. It did not apply to thermal images, which have different characteristic regions resulting from internal and external heat sources.

The revised signature metric still requires that the user designate where the object is for which he wants the metric computed. The model does not find the target outline. The target outline includes the target's shadow. In the test analysis, the target mask covered the entire projection of the vehicle onto the plane, including those parts of the scene that were foreground obscuring portions of the target. Further research is needed to address whether or not foreground obstruction should be excluded.

The revised signature metric computation begins by searching for the appropriate perceptual organization of the target. The target is organized into either (1) a single region (the target silhouette), or (2) two regions, nominally corresponding to the illuminated and shadowed regions (Two-tone representations are sufficient for 3D shape/structure perception [6]). The organization finds the sub-region of the target with the largest value of a region distinctiveness metric, and its complement.

The algorithm divides the target into light and dark regions via a threshold on the achromatic visual channel. It searches for the threshold that produces the primary target region (either above or below the threshold) with the highest region distinctiveness metric. The theory is that pre-attentive visual processing organizes the scene into distinctive shapes based on individual region distinctiveness, not on the combined effects of adjacent regions; highly distinctive regions are perceived first, and lower distinctiveness regions later. The perceptual organization consists of the primary target region (which may be the entire target projection) and its complement (which may be null).

Figure 5 illustrates the results of the perceptual organization algorithm. It shows the original image, the target mask, and the 2-region perceptual organization. Were it not for the gun tube, the target mask alone would look like a blob not a vehicle, but the 2-tone perceptual organization has the look of a vehicle.

Fig. 5: Example Image, Target Mask, and 2-Region Perceptual Organization

The processing flow to compute the region distinctiveness metric, RDM(.), is essentially the same as that used in the previous VDM2000 model [1]. There was no change in the luminance and color appearance transform from RGB to Acd (achromatic and color-opponent) coordinates.

The vehicle signature metric combines the region distinctiveness metric with a measure of the magnitude of shape information in the component regions. The measure of the shape information of a region is based on the dispersion of the region, i.e., the variance in the location of the pixels in the region (or, equivalently, the mean squared distance from the center of mass):

$$\mathrm{Dsp}(\mathcal{R}) = \Sigma_{\mathrm{rc\ in}\ \mathcal{R}} [(r - \mu_r)^2 + (c - \mu_c)^2] / \mathrm{Count}(\mathcal{R}) \quad (3)$$

The dispersion divided by the area is a measure of the compactness of the region. A commonly used alternative measure of the shape complexity or compactness of a region is the perimeter squared divided by the area. However this simple metric has some serious shortcomings: a shape that is almost a circle but with ragged edges can have a very large value, and the metric for a shape consisting of disjoint regions is independent of the configuration of the disjoint regions. The dispersion metric does not have these problems.

The vehicle signature metric is equal to the sum of the region distinctiveness metrics, times the sum of the region dispersions, Dsp(.), divided by the sum of the region areas, Area(.) for the two perceptual regions $\mathcal{R}1$ and $\mathcal{R}2$. The metric for the vehicle is based on the combination of the complementary regions:

$$M = (\mathrm{RDM}(\mathcal{R}1) + \mathrm{RDM}(\mathcal{R}2)) * (\mathrm{Dsp}(\mathcal{R}1) + \mathrm{Dsp}(\mathcal{R}2)) / (\mathrm{Area}(\mathcal{R}1) + \mathrm{Area}(\mathcal{R}2)) \quad (4)$$

The psychometric function is unchanged from the earlier model formulation. The entire model has only one free parameter: the vehicle metric value for 50% probability of detection.

Figure 6 shows a scatter plot comparing the model prediction of Pd with the empirical estimate of Pd for all 800 images with targets in the cued detection experiment. Figure 7 compares Pd in the cued detection experiment with Pd from the search experiment. Figure 8 compares the model prediction of Pd in search with Pd from the search experiment. The model accounts for 72% of the variance in Pd in the cued detection experiment. Pd in the cued detection experiment accounts for 81% of the variance in Pd in the search experiment. The chain rule predicts that the model should account for 58% of the variance in Pd in the search experiment. In fact, the model accounts for 61% of the variance in Pd in the search experiment.

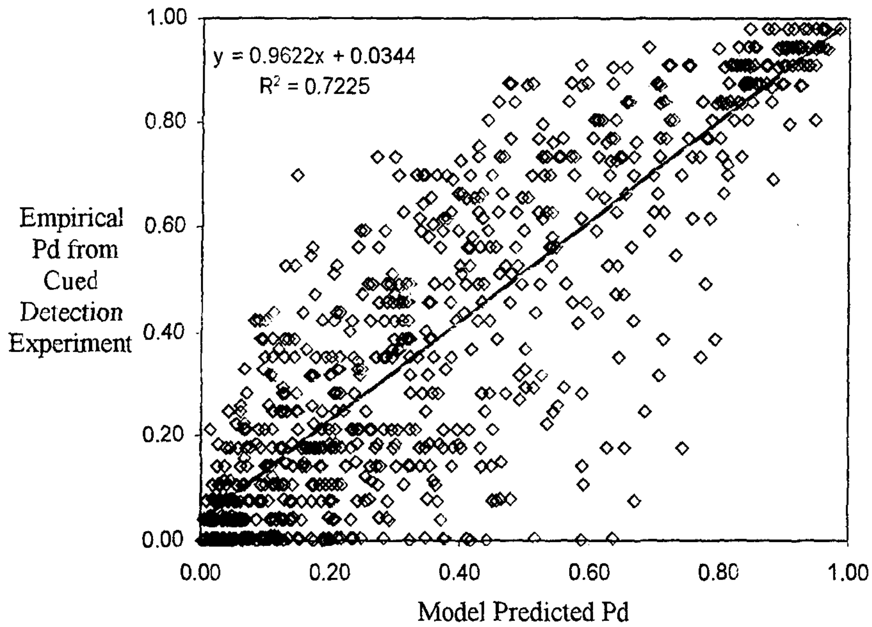


Fig. 6: Comparison of Model and Empirical Pd for Cued Detection Experiment

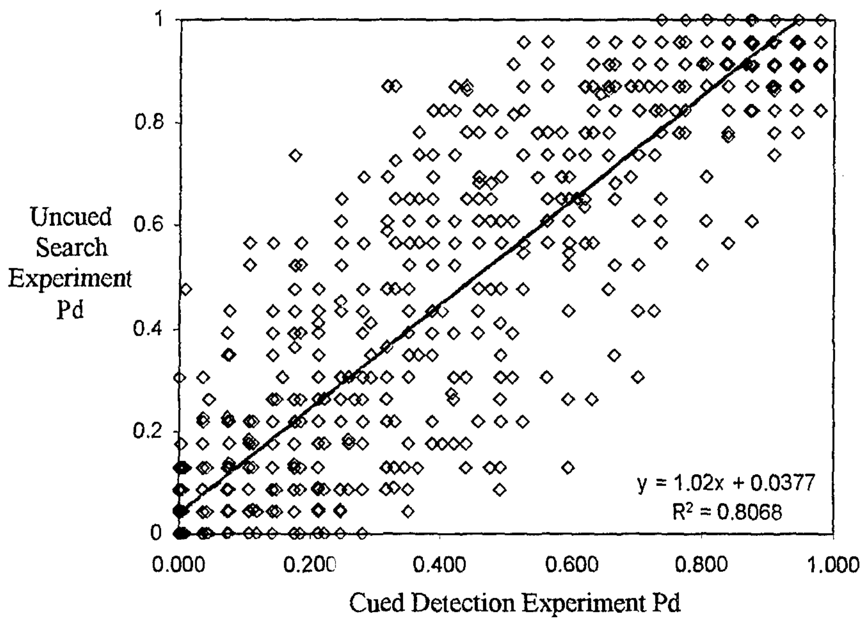


Fig. 7: Comparison of Empirical Pd for Cued Detection and Uncued Search Experiments

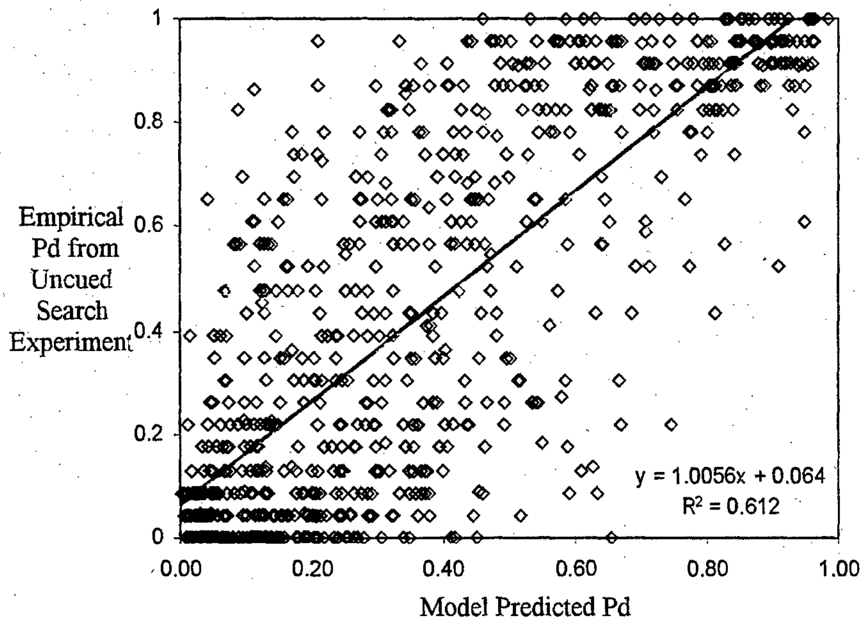


Fig. 8: Comparison of Model and Empirical Pd for Uncued Search Experiment

5 Discussion

The proposed model is derived from a theory of search behavior assembled from empirical and theoretical research in human vision and decision making, and is substantiated by comparison to the results of large-sample search and target detection experiments.

The proposed model is similar in structure, but differs in specifics, from the widely-used U. S. Army NVESD search model [7]. Both models predict the probability of detection within some specified time as the product of the probability of detection and the probability of search completion within the specified time. The proposed model uses a lognormal formulation for the distribution of search time, whereas the NVESD model uses a negative exponential formulation. The proposed model uses a constant for one parameter of the lognormal distribution, and calculates the other parameter as a linear function of the probability of detection. The NVESD model calculates the rate parameter of the negative exponential distribution as proportional to the probability of detection. The proposed model uses a simple, one parameter psychometric function. The NVESD psychometric function uses three parameters. The greatest difference between the two models is in the vehicle signature metric. The NVESD metric is a function of the target area, target-to-background contrast and range. The proposed metric is an image-based computational vision formulation that accounts for target size, shape, perceptual organization, local clutter, local color and luminance contrast across the target boundary.

The vehicle signature metric in the earlier VDM2000 model was not applicable to thermal images since it employed a simple 3D geometry appropriate to solar illumination and reflected signature. It was not applicable to long-wave infra-red images, in which the perceptual regions correspond to hot and cold physical components. The revised vehicle signature metric discovers the appropriate perceptual organization. It is potentially applicable to thermal images, but has not been tested.

The refined signature metric still has one important limitation. The problem occurs when an edge of the target is in line with a linear feature in the scene, and when the contrast across that target edge is close to the contrast across the background feature. When this occurs, human visual perception tends to see the target edge as a continuation of the background feature. When the contrast difference is low, the perceptual organization that segments the background feature dominates the segmentation of the target, and probability of detection is low. The computational method has no way to detect this condition, and so

overestimates the probability of detection. This condition occurred in approximately 100 of the 800 images used in the perception experiments.

Acknowledgments

This research was funded by the U.S. Army TACOM under Small Business Innovation Research contract DAAE07-97-C-X101.

References

1. Witus, G., Karlsen, R. and Gerhart, G. 2001. The VDM2000 visual vehicle detection model – theory and validation. *Proceedings of the 12th Annual Ground Target Modeling and Validation Conference*, Houghton, MI. 260-268.

2. Cartier, J. S. and Hsu, D. H. 1995. Human visual search: a two state process. *Proc. SPIE Vol. 2470*, p. 58-68, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing VI, Gerald C. Holst; Ed.

3. Ulrich, R. and Miller, J. 1993. Information processing models generating lognormally distributed response times. *Journal of Mathematical Psychology*. 37, 513-525.

4. Shapiro, S. S., and Wilk, M. B. 1965. An analysis of variance test for normality. *Biometrika*. 52, 591.

5. Wolfe, J. M. 1994. Guided Search 2.0: a revised model of visual search. *Psychonomic Bulletin and Review*. 1, 202-238.

6. Moore, C., and Cavanagh, P. 1998. Recovery of 3D volume from 2-tone images of novel objects. *Cognition*. 67:45-71.

7. O'Kane, B. L. 1995. Validation of prediction models for target Acquisition with electro-optical sensors. In *Vision Models for Target Detection and Recognition*, E. Peli, Ed. World Scientific Press. 192-218.

An Investigation of the Reliability of Search Statistics
Based on Results from Paired Images

James R. McManamey
U.S. Army, Night Vision and Electronic Sensors Directorate
Fort Belvoir, VA 22060-5806

Abstract

For more than 50 years, the Department of Defense and its contractors have been conducting search experiments in the field and with actual and simulated field imagery directed toward development of observer and sensor performance models and psychophysical evaluation of camouflage and signatures of military assets. Model critics have pointed to lack of agreement between model predictions and experimental results, seeking explanations for a perceived lack of correlation. This paper exploits data from a perception experiment to show what can be expected in regard to the consistency of psychophysical quantities such as response time, time for a correct response, average time for a "no target" response, and probability of detection. It examines the difference in consistency between averaged values and raw values from individual subjects. It also examines the difference between average time for a correct response versus the average time for any response (correct or incorrect), as well as the use of average versus median values. The methodology involves the use of paired images. While most of the images in the experiment were shown to the observers only once, a few of the images were shown twice to each observer. The second time the images were shown, they were flipped horizontally (that is, a mirror image was used with a reflection about the vertical axis). Some critics have claimed that, under such circumstances, the observer "learns the image" the first time it is presented. The hypothesized result is that the second time the image is seen, the observer will respond more quickly and more accurately. Analysis is shown indicating that, at least in these experiments, no demonstratable learning has taken place. Analysis also shows that individual (raw) response times are not very predictive of a second observation of the same image by the same individual ($r^2 \approx 0.3$) but averages over all individuals for one image are highly predictive of a second observation of the same image by the group of observers ($r^2 > 0.9$). It shows that the improvement is even more dramatic when response times are restricted to correct responses only. Finally, it shows that probability of detection is a relatively consistent statistic for an image and its mirror image ($r^2 > 0.95$).

1. Background

Visual search, whether pursued with the unaided eye, aided by binoculars or other optical systems, or employing any of an increasing diversity of electro-optical devices, has long been an activity of high military importance. For centuries, sentinels occupying elevated observation posts could provide warning of an enemy's approach in time that defenders could be readied. On the modern battlefield, early warning, reconnaissance, and target acquisition remain very important applications of search. Thus, it is not surprising that the Department of Defense (DoD) and its various contractors began more than 50 years ago to model various aspects of the visual search process. One pioneer was H. Richard Blackwell who, during the years 1946 to 1951, conducted a series of experiments for the office of Naval Research[1]. Since that time, countless numbers of investigators have conducted similarly uncountable experiments, both in the field and in the laboratory, in order to evaluate and/or better predict the performance of observers, sensors, decoys, and camouflage.

Depending on what observer performance model one uses, the nature of the target and background, and the predictions one tries to make, it is not uncommon to find substantial discrepancies between measured and predicted results. In a 1992 paper discussing such discrepancies, Nichols and Paik identify two kinds of anomalies that are major contributors to such discrepancies, illustrating their point with a graph upon which figure 1 is based. They refer to these as "contrast anomalies" due to such factors as shadows and highlights on the target, and "clutter effects," especially "high subjective clutter," which

is generally presumed to be a characteristic of the scene or that part of the scene in which the target is located. While they do not indicate how typical the data in figure 1 might be, nor the exact experiment upon which it is based, they represent it as a common occurrence[2]. Indeed, such results are commonly seen within the community and supported by such results as those reported by D'Agostino[3]. In an experiment performed as part of a Small Business Inovative Research (SIBR) contract, Witus reported modeling improvements that result in substantially less scatter than that depicted by Nichols and Paik. Nevertheless, the Pearson Correlation Coefficients typically obtained by Witus ranged from about 0.80 to about 0.87, which might be considered low by those who are unaware of typical results in this field of endeavor (see figure 2), but in actually these results are quite good[4] (compare to Nichols and Paik in figure 1 and to D'Agostino).

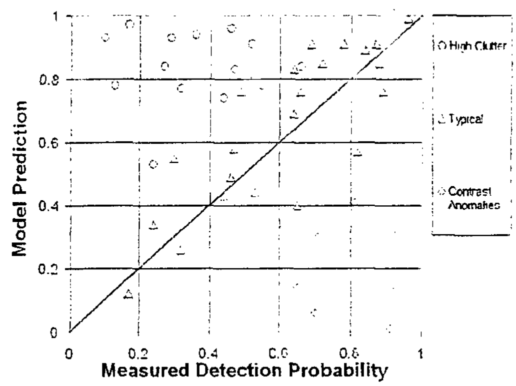


Figure 1 – Adapted from Nichols and Paik

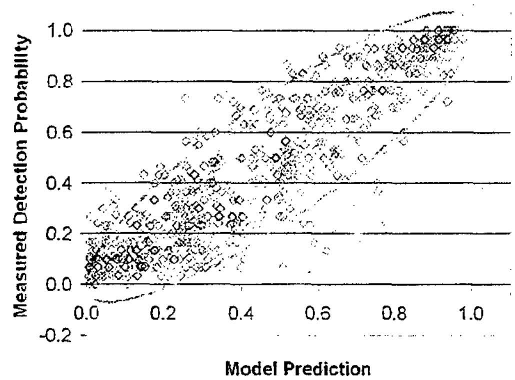


Figure 2 – From Witus

As in the examples above, it is common for those modeling visual search to ask how well the model predicts actual human behavior, but it is not common to ask how consistent that behavior actually is. Witus reported that, at best, 25 percent of the variance between model and observation remained unaccounted for. Some might be critical, speculating that the model must be failing to account for major sources of variance, but what if the remaining major source of variance is the observer himself? It is possible that the most significant variable remaining can be thought of as in internal state of the observer that, for our purposes at this time, is a random variable. If this is so, it is important to know the distribution of this variable so that suitable bounds can be placed on model predictions and so that we do not waste our time seeking additional external quantities to account for variations in experimental results. This paper reports analysis of a portion of the data from a set of experiments run by the U.S. Army Night Vision and Electronic Sensors Directorate at Fort Belvoir, Virginia. It was a field-of-view search experiment using synthetic thermal images. The experiment is known as Search Experiment A.

In Search Experiment A, trained military observers were systematically conditioned to respond quickly and accurately to field-of-view search stimuli. Two hundred different stimuli were used, however there were 16 paired stimuli that are of special interest for the analysis reported here. Each pair consisted of two images that were different only in one being the reverse (mirror image) of the other (i.e. the image was flipped, left to right). The order of presentation was randomized so that about half of the time one of the paired images was presented first and the other half of the time the corresponding flipped image was presented first. The number of other images presented between the paired images was also a random variable. Thus, these paired image presentations were independent tasks of equal difficulty and data from these items could be used to test for reliability of search statistics (that is, their repeatability). The following observations are based on analysis of these 16 stimuli in this field-of-view search task with a 12-second time limit.

There were 26 observers in the data set. All of the images in this analysis contained a target, although some of the other images in the experiment contained no target. For each stimulus, each observer had 3 response options:

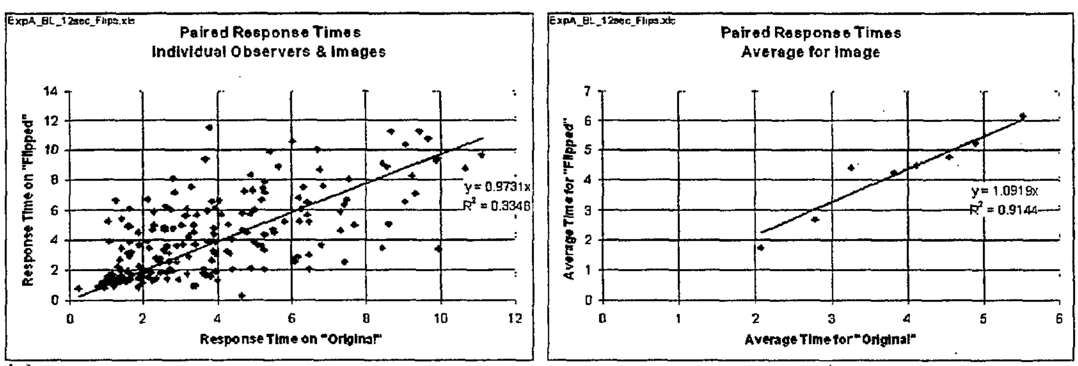
(1) indicate the position of something believed to be a target,
(2) indicate that there was nothing that the observer believed was a target, or
(3) fail to respond within the 12-second time limit.

Responses of the first type may further be classified as correct responses or false alarms, depending on whether or not they clicked within the scoring box of a target. Responses of the second type may further be classified as missed targets or correct null trials, depending on whether or not there actually was a target present. All of the images in this analysis contain a target, making "no target" responses universally incorrect. Even so, 31 percent of the images in the full dataset lacked a target, and the observers answered "no target" 23 percent of the time overall for the 200 stimuli. Thus, "no target" was a reasonable expectation on the part of the observer. Since each of the 8 unique images was presented twice, and since there were 26 observers, there are 416 opportunities to respond in this analysis. In 8 instances, an observer failed to respond within the time limit. Two observers failed to respond twice, and 4 of them failed to respond once. The remaining 20 observers responded to every item within the time limit.

2. Response Time

Response time is the time elapsed from the presentation of the stimulus until the observer responds. Correctness of the response is not considered in determining response time. However, failure to respond is not considered a response, so timeouts are omitted. Because timeouts account for less than 2% of all opportunities to respond, they are ignored. The values used here are based on raw response time, with no correction for interface induced delays or the effect of timeouts.

As indicated in section 1, there were 416 opportunities to respond (208 pairs), but there were only 408 actual responses. This is because there were 8 instances in which the observer failed to respond within the 12-second time limit. In every instance of a timeout, it occurred for only one of the two presentations of the image in question. When timeouts are omitted, it is also necessary to omit the responses that would be paired with them. Thus, there remain a total of 400 responses (200 pairs). If we plot the response time for each observer for each "original" image against the corresponding response time for the "flipped" image, we obtain figure 3-a. We see that there is a weak correlation between the two response times. Only 33% of the variance is explained by the assumption that the response times should be equal. However, when the individual response



(a) (b)

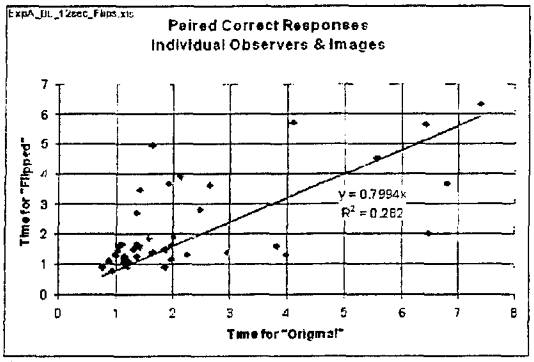
Figure 3 -- Stability of the response time (without regard to correctness of the response) in a search experiment. Individual paired observations (a) and averages for each image (b) are shown for 8 images each shown to each observer twice

times are averaged so that we obtain an average response time for each image, the result is figure 3-b. This shows that 91% of the variance is accounted for by the assumption that the average response time for each original is equal to the average response time for the corresponding flipped image. Thus, we observe that, while response time varies substantially due to various unidentified factors, the average response time for an image is relatively stable.

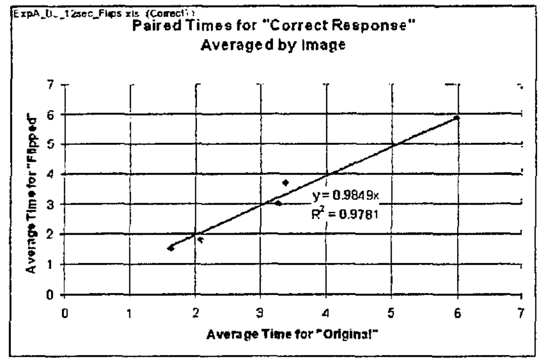
3. Average Time for Correct Response

Incorrect responses may occur for many different reasons. In addition to targets that are hard to find and false alarms that are particularly target-like, incorrect responses may be due to a host of other causes such as lack of diligence on the part of the observer and accidental responses. By their nature, we suspect that the response time for incorrect responses is likely to be

more variable than for correct responses. Thus, one might well ask what happens when incorrect responses are filtered from the results above. While there were 416 opportunities to respond (208 pairs), as indicated above, there were only 408 actual responses after omitting timeouts. In addition, in 109 instances, the observer indicated there was no target and in 179 instances they "false alarmed," leaving only 120 correct responses. Finally, in 34 instances, an observer had a correct response for only one of the two image presentations in a pair. This leaves 86 paired correct responses (43 pairs). In this selection process, 5 of the 8 images survive the selection criteria with at least 3 paired responses each. In figure 4, these paired responses are shown both individually (a) and with the time averaged by image over the observers (b).



(a)



(b)

Figure 4 – Stability of the time for a correct response in a search experiment. Individual paired correct response times (a) and average time for correct responses (b) are shown for 6 images each presented to observers twice.

One will immediately observe that only 28% of the variance is explained by the pairing of individual response times (left), which is worse than in the unfiltered case in section 2. This indicates that, as is the case with the raw response time, the time it takes an individual for a correct response to an image is highly variable and not a good predictor of future response time by the same individual on an equivalent search task. Again, it is evidently affected by a number of unidentified factors. One will also observe that 98% of the variance is explained by the pairing of average times for correct responses (figure 4-b). This indicates that the average time it takes for a group of individuals to give a correct response to an image may be highly consistent. Note that the range of averages is from about 1.5 seconds up to about 6 seconds. Thus the range is about 3 times the minimum value, indicating that there were both easy and difficult search tasks represented.

4. Average Response Time versus Average Time for a Correct Response

Based on sections 2 and 3 above, we see that average time for a correct response is more reliable (consistent) than the average time to respond when response correctness is not considered. Thus, one might well ask to what extent these two quantities measure the same thing. If they are actually different estimates of the same quantity, it seems the more reliable of the two (average time for a correct response) should be preferred in all cases, but that the two can be used interchangeably when, for some reason, one is available and the other is not. On the other hand, if they are not separate estimates of the same quantity, the investigator needs to be especially careful to determine which is required in a particular situation and not confuse the two. In fact, they do not appear to actually measure the same thing although there is a relatively weak correlation between them ($r^2 = 0.64$, see figure 5). Also notice that, in the cases involved here, the average time for a correct response is only 85% of the average time for all responses of the average time for all responses to the same image. In other words, those observers who respond most quickly are generally also the most accurate. Although this phenomenon has been observed before, in this case the results are not statistically significant[5].

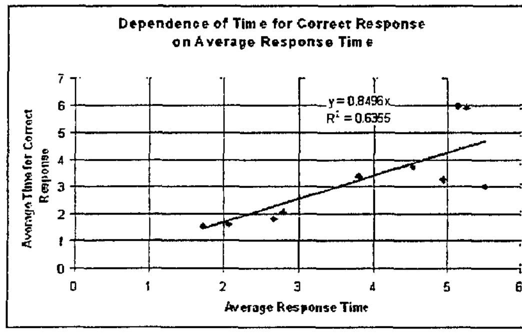


Figure 5 -- Relationship between average response time and average time for a correct response to the same image.

When the one-tailed sign test is applied using a 0.05 confidence level, the null hypothesis,

$$H_0 : \bar{t}_c = \bar{t}_a , \tag{1}$$

where $\bar{t}_c$ and $\bar{t}_a$ are the average time for correct responses and the average time for all responses respectively, cannot be rejected because there are only 8 of the 10 pairs in figure 5 for which $\bar{t}_c < \bar{t}_a$. At the 0.05 confidence level, 9 out of 10 is required to reject the null hypothesis.

While the relationships observed in the above paragraph between $\bar{t}_c$ and $\bar{t}_a$ are not statistically significant, related more general observations are highly significant. Overall in the Search A experiment with the 12-second time limit, the average response time for all individuals to all items with timeouts omitted ($\bar{t}_a$) is 4.0959 seconds with standard deviation of 2.6158 (5069 instances). The subset of these responses which were correct have an average $\bar{t}_c$ = 2.5310 seconds with standard deviation of 2.5222 (1943 instances). If we believe that $\bar{t}_c < \bar{t}_a$, the null hypothesis that must be rejected to show statistical significance is the same as stated above. Since our hypothesis states the direction of the inequality, we use a one-tailed test. We obtain –9.8725 for the *t*-statistic (Student's *t*-distribution). With nearly infinite degrees of freedom, this permits rejection of the null hypothesis at virtually any level of significance one may desire, indicating extremely high statistical significance. Similarly, restricting our consideration to the images that were used twice, $\bar{t}_a$ =4.1326 seconds with standard deviation of 2.6563 (408 instances), while $\bar{t}_c$ = 2.9249 seconds with standard deviation of 2.3192 (120 instances). In this case, the value of the *t*-statistic is –5.7046. Again, we have nearly infinite degrees of freedom and the null hypothesis can be rejected with extremely high confidence. In either case, we come to the inescapable conclusion that the average time for a correct response is significantly less than the average for all responses.

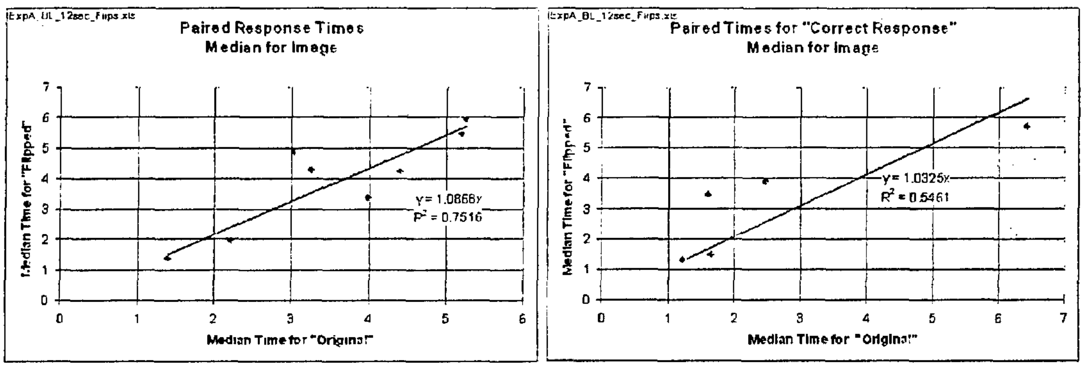
Observations that have been largely unconfirmed statistically, indicate that quick responses are most commonly associated with correct detections when targets are present and with false alarms when targets are not present. On the other hand, slower responses are most commonly associated with false alarms when targets are present, correct "no target" decisions, and increased probability of timeouts. Furthermore, different observers have different mixtures of these response characteristics but tend to be fast, medium, or slow. As such, fast observers tend to have more items correct when targets are present and more items incorrect when they are not, while slow observers tend to have more items correct when targets are absent and a higher percentage of missed targets (either responding "no target" or false alarming). In the end, with a good cross-section of observers and an experiment that is well balanced between easy, medium, and hard detections and a reasonable number of no-target (null) scenes, these effects tend to cancel each other out to some extent. Assuming these observations to be correct, average response time would be an effect characteristic of the observers (fast, medium, or slow), while average time for a correct response would be an effect more characteristic of the image set (difficulty and proportion of nulls).

The conjectures stated in the preceding paragraph cannot be tested on the limited set of items being analyzed here. In the case of the experiment as a whole, there may be sufficient data to check some aspects of these conjectures, but a full test would probably require a data collection specifically designed for that purpose.

5. Average versus Median Response Time

In figures 3-a and 4-a above, we observe that there appears to be a higher density of short response times than of longer response times. Indeed, this is born out by the fact that the mean response time is 4.0 seconds, while the median is 3.4 seconds and the mode is only 1.6 seconds. Furthermore the skewness of the distribution is 0.89. Thus, it may be that the median response time would be a better indicator of central tendency than the mean. For this reason, the correlations using the median were compared to the correlations obtained in sections 2 and 3 above. Figure 6 shows the result.

One will observe that in both of these instances, the correlation between the median values is lower than the correlation for the average values (figures 3-b and 4-b). In the case of time for correct responses, the assumption of equality of medians accounts for only 54 percent of the variance, whereas for averages it is 98 percent. For response time without regard to the correctness of the response, the medians account for 75 percent of the variance compared to 91 percent for the averages. There is no explanation offered here for this unexpected result. However, this phenomenon was investigated further for "no



(a) **(b)**

Figure 6 – Median Response Time (a) and Median Time for Correct Response (b) for Paired Images

target" responses (see section 7 below) and overall response time for observers (see section 8 below). In one of these cases (observer overall response time), the median continued to be less consistent than the average, explaining 64 percent of the variance compared to 79 percent for the average. However, in the other case (time for "no target" responses) the median was better than the average, accounting for 85 percent of the variance compared to 56 percent for the average.

6. Response Time for First Presentation versus Second Presentation

6.1 *Systematic Experimental Error due to Learning*

It has already been pointed out that the correlations between paired responses improve when average response time is used (as in figures 3-b and 4-b) rather than individual response times (as in figures 3-a and 4-a). However, there arises the issue of why this is so. Certainly, with any variable, a single measurement is prone to some measurement error. As suggested by the central limit theorem of statistics, repeating the measurement and averaging the values obtained can substantially improve the reliability of the measurement by reducing measurement error. However, in at least one sense, we are dealing with another phenomenon here. We actually believe some observers are different than others and expect a distribution of response times for a given image. We also expect the same observer to be different from one time to the next as the result of various psychophysical phenomena. How much sleep did she have last night? Did he get cut off in traffic this morning? These and a thousand other things can affect the observer's performance from day to day and from moment to moment. Thus, the variability that is observed in figure 3-a is the result of numerous factors, sometimes adding to and sometimes subtracting from the individual observer's response time. In addition to measurement error, there are other plausible sources of the scatter observed in figures 3-a and 4-a. Some of these sources are, like measurement error, random. However, systematic experimental errors are of particular concern. The balance of this section investigates the possibility of one source of systematic experimental error.

Some in the community have suggested that, in some cases where images are repeated, observers "learn the images in a test set." If this happens, it can be expected to produce systematic experimental errors. One could readily argue that, in the analysis above, learning has produced a significant portion of the effect observed between figures 3-a and 3-b. If an observer learned an image the first time they saw it, they could be expected to produce the same response the second time, but more quickly. Since the images were presented in a randomized order, they would sometimes respond more quickly to the "original image" because they had seen the "flipped image" first, and the rest of the time the reverse would be true. The net effect would be to produce "random scatter" in (figures 3-a and 4-a), but that this effect would be "averaged out" when results are pooled by image (figures 3-b and 4-b). If such learning has occurred here, we should avoid using repeated occurrences of images in the future and should avoid placing too much credence to conclusions drawn from this analysis of repeated images. However, if such supposed learning has not significantly influenced the search time for these images, it may enhance the value of including such analyses in future experiments. While the designers of this experiment thought that other factors negated the likelihood of any significant learning effect, it is best to test the conjecture.

6.2 *Intuitive Test for Learning*

In figure 7-a, we see the results of plotting the same data as that in figure 3-a, with one change. Rather than have the response to the original image on the x-axis and the flipped image on the y-axis, the x-axis represents the response time for the first of the two images to be presented, whether that was the original or the flipped image, while the y-axis represents the response time for the second image to be presented. If in actuality the observers needed less time to respond the second time they saw the image, with the data presented in this way, one would expect a higher correlation than was obtained in figure 3-a. One will observe that there is negligible change in the correlation.

Similarly, in figure 7-b, we see the results of replotting the same data as that in figure 4-a, again with a single change. As in figure 4-a, only correct responses are presented, but rather than have the response to the original image on the x-axis and the flipped image on the y-axis, the x-axis represents the response time for the first of the two images to be presented and the y-axis represents the response time for the second image to be presented. This is the situation in which one would expect learning would have its greatest impact on the results. Specifically, this represents cases in which the observer was correct both times and would be in the best position to utilize what they had learned the first time they saw the image. In this case, the correlation has increased from figure 4-a ($r^2 = 0.28$) to figure 7-b ($r^2 = 0.36$). The question then becomes a matter of the significance of this increase. We contend that this is not a significant increase. The following explains why.

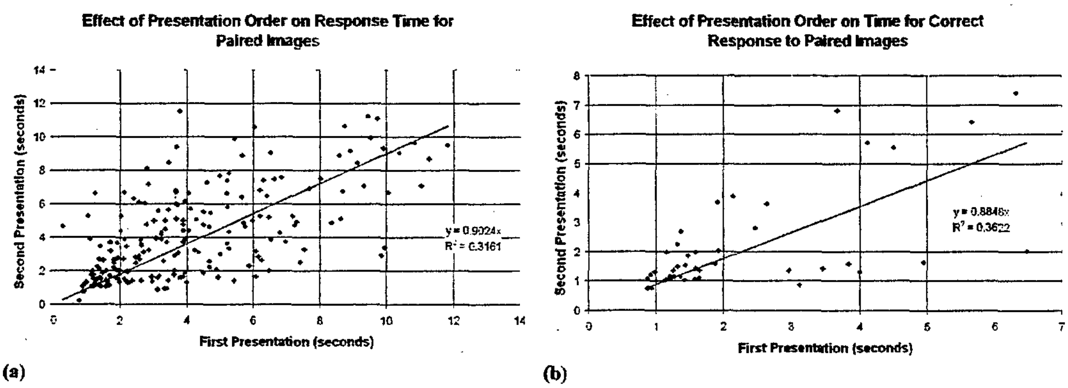


Figure 7 – (a) Observer response time for the first presentation of an image compared to that for the second presentation of the image. (b) Same as (a) except correct responses only.

It might be proposed that figures 3-a and 7-a give two estimates of the correlation to be expected in 4-a and 7-b. In fact, figure 3-a ($r^2 = 0.33$) gives almost exactly the same value as figure 7-a ($r^2 = 0.32$). Because the sample is smaller in figures 4-a and 7-b, we can expect more variability. The results in these two cases are on either side of the results in figures 3-a and 7-a, with only 0.05 separating 3-a and 4-a and only 0.04 separating 7-a and 7-b. Since these changes are of about the same magnitude, and since the change from 7-a to 7-b is actually less, we suspect that the increase in correlation is not significant.

6.3 *Using the t-test to Check for Learning*

There are 41 instances in which an individual responded correctly to both an image and its flipped variant. Thus, for these instances, there are 82 response times, one the first time each image was shown to an individual and a second one the second time each image was shown to the same individual. The mean time for these 82 responses is $\mu = 2.2864$ and their standard deviation is $\sigma = 1.6394$. The average time taken to respond in those 41 cases where an image was shown a second time is $m = 2.2374$ and the standard deviation of the sample is $s = 1.7489$. Our hypothesis states

$$H_1 : m < \mu . \tag{2}$$

In the strictest sense, this is true, but it could be a result of random fluctuations due to sampling. Thus, the null hypothesis, which must be rejected to accept the hypothesis as statistically significant, states

$$H_0 : m \geq \mu . \tag{3}$$

We wish to test the hypothesis at the 90% confidence level. That is, if we cannot be at least 90% confident that H_0 is in error, we will accept it. We will apply a t-test. Using standard references, $-1.3031 \leq t_{0.10} \leq 1.3031$ is the critical region for the 40 degrees of freedom in this case. If the statistic t lies in this region, we will accept the null hypothesis. We use the formula

$$t = \frac{m - \mu}{s / \sqrt{n}} . \tag{4}$$

Thus,

$$\begin{aligned} t &= \frac{2.2374 - 2.2864}{1.7489/\sqrt{41}} \\ &= -0.1794 \end{aligned}$$

Since t lies well within the critical region, we cannot reject H_0. Thus, the average time to correctly respond to these items the second time is not significantly less than the first time. In fact, this value of t indicates that there is only about a 0.57 probability that m is really less than μ.

6.4 *Using the Sign Test to Check for Learning*

The t-test makes the assumption that either the response times are normally distributed or the sample averages are approximately normally distributed. Certainly, the response times are not normally distributed (see section 5 above). Since we cannot be sure the sample averages are close enough to normally distributed, it is wise to confirm the above results using a distribution independent test. The sign test is suitable in this case.

When each response time for the second presentation of each image to an individual is subtracted from the corresponding response time for the first presentation of that image, there are 20 positive differences and 21 negative differences. This indicates that the time for a correct response the second time an image was seen was less than the time for a correct response the first time it was seen (supporting H_1) in 20 of the 41 instances. For a binomial distribution where there are n = 41 instances of two mutually exclusive outcomes, each with a probability of ½ (as would be the case under H_0), the probability of having the less frequent outcome occur 20 times or less is 0.50. It would need to occur less than 16 times (i.e. 15 time or less) before the probability of H_0 being correct would be less than 0.10. Thus, under the sign test we cannot reject H_0 with 90% confidence, so we do not accept H_1.

6.5 *Conclusions Regarding Learning*

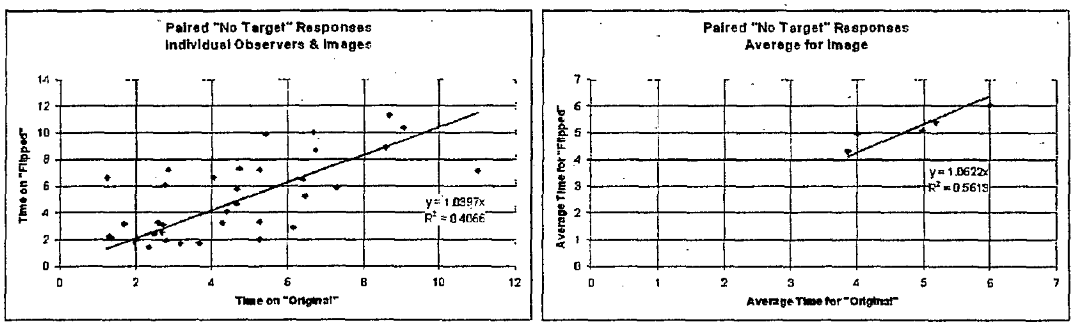
Response times have been examined for the first and second presentation of several images. They have been compared intuitively, using the parametric t-test, and using the non-parametric sign test. All three methods lead to the same conclusion: no demonstratable learning has taken place during the test. Thus, the use of images a second time after flipping should not bias the recorded response times.

7. Average Time for "No Target" Response

There is considerable interest in the question of how long an individual takes to decide that there is no target in a scene. It has already been shown elsewhere that there is no significant difference in the way an observer behaves when they cannot find the target than in a situation where there is no target[6]. Thus, although there was always a target present in these images, it makes sense to examine those cases of an individual subject deciding that there was no target, and treat them as "no target" trials. As indicated in section 3, there were 109 such instances in the 416 opportunities to respond within this data subset. However, in 37 of these 109 instances, the individual responded "no target" to one of the paired image presentations and either found the target, false alarmed, or timed out on the other image presentation. Thus, 72 responses (36 pairs) survive for

analysis. These represent from 4 to 12 responses for each of 5 images, there being 3 images for which no subject responded "no target" for both image presentations.

Again, we observe that the individual raw response time on one trial is not a good predictor of the individual raw response time on another trial. Only 41% of the variance is explained by the assumption that these two quantities should be equal. However, observe that this is higher than the 28% obtained for average time for a correct response or the 33% obtained for average response time. Even so, the reliability of the average for all observers over an image (figure 8-b) is much lower than for all responses (figure 3-b) or for correct responses (figure 4-b). Some of this may be due to the smaller sample size for "no target responses," but this hypothesis has not been tested for statistical significance.



(a) (b)

Figure 8 -- Paired "no target" response times. Individual responses are shown on the left (a) and responses averaged by image are shown on the right (b).

8. Average Response Time per Subject

There is evidence to suggest that an individual response time for an image is a function of at least two variables. These are the observer's predisposition to be fast or slow and the difficulty of the detection task represented by the image. Thus, in addition to the average response time for all observers to a single image, we may be interested in how consistent each individual observer's time is when averaged over several images. The raw data for figure 9 is the same as figure 3. However, whereas figure 3-b represents averages for images over a set of observers, the graph in figure 9 represents averages for observers over a set of images. A substantial portion of the increased variability in figure 9 compared to the figure 3-b is probably due to the smaller sample size. In figure 3-b, each point is an estimate of the mean based on 24 to 26 observations (the number of observers who responded within the 12 second time limit), while in figure 9, each point is an estimate of the mean based on only 6 to 8 observations (the number of images to which the observer responded). One will notice that the assumption that an observer's average time over the set of images on one trial predicts their average time on a second trial accounts for about 79 percent of the variance.

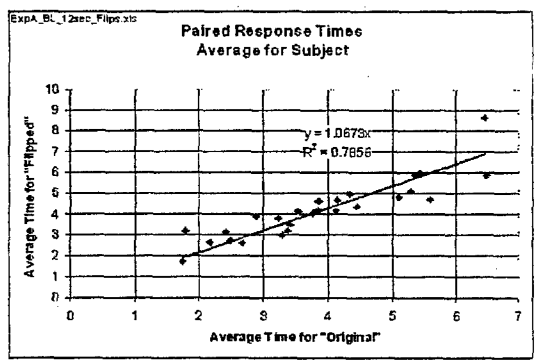


Figure 9 -- Average response time for observers on paired sets of equivalent images.

9. Probability of Detection

In addition to the time to detect the target in a field-of-view and the time to decide that there is no target in a field-of-view, we are frequently interested in knowing the probability of detecting the target. The values shown in the graph (figure 10) are based on the 416 opportunities to respond that are described in sections 1, 2, and 3 above. A detection is indicated by clicking within the permitted distance of the target within the 12-second time limit. In this case, a timeout is considered a failure to detect within the 12-second time limit. Thus, the sample size in each case is 26 observers. The probability of

detection is the number of observers who detected the target divided by 26 (the total number of observers). The assumption that the probability of detecting the target in the flipped image is the same as that for the original image, accounts for 95 percent of the variance. If the probability of detection is replaced by the conditional probability of detecting the target given a response, the results are nearly identical.

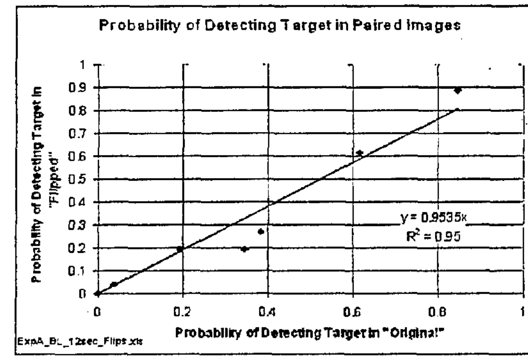


Figure 10 -- Probability of detection for paired images.

10. Conclusions

The response time for individual observers to individual images is highly variable whether or not incorrect responses are included in the calculation. In such cases, an observer's time to respond to an item in one instance is not a reliable predictor of their time to respond in another instance, even for the same image. However, when the response time for a number of individuals is averaged, this average is a reliable statistic for a given image, especially if consideration is limited to correct responses. Even though the distribution of response times for a set of images may be skewed, this analysis has indicated that average response time is a more reliable statistic than median response time. This analysis has also indicated that the time for a "no target" response is not a very reliable image statistic, but the probability of detecting the target in an image with a single target is a reliable statistic. Finally, different observers respond at different rates, but an individual observer's average response time is a relatively reliable statistic.

Given the reliability of probability of detection estimates from the 26 observers in this analysis, it seems likely that 8% to 15% of the unexplained variance observed by Witus is due to factors external to his observers. Thus, those seeking to further improve observer performance models can reasonably expect to find additional relevant parameters.

The conclusions reached here must be considered to be tentative. However, in future search experiments, it might be a good idea to include more pairs of identical or flipped images. Replication of these results with other image sets would do much to establish the generality of these conclusions.

11. References

[1] H.R. Blackwell, *Psychophysical Thresholds: Experimental Studies of Methods of Measurement*, Engineering Research Bulletin No. 36, University of Michigan Press, Ann Arbor (1953).

[2] W. Nichols and H. Paik, "A Methodology for Evaluating Clutter Effects on Observer Detection Performance," *Proceedings of the Third Annual Ground Target Modeling and Validation Conference, Keweenaw Research Center*, Houghton, MI (1993).

[3] J. D'Agostino, W. Lawson, and D. Wilson, "Concepts for search and detection model improvements," *Proc. SPIE Vol. 3063, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing VIII*, ed. G. C. Holst (1997).

[4] U.S. Army contract DAAE07-97-C-X024.

[5] Barbara O'Kane, U.S. Army, Night Vision and Electronic Sensors Directorate.

[6] Stephen F. Sousk, U.S. Army, Night Vision and Electronic Sensors Directorate, Private Communication about 15 January 2001.

Search Performance With PTN Successfully Predicts Search Performance With An Advanced Sensor

Melvin Friedman
Timothy C. Edwards
Max Lorenzo
Richard Vollmerhausen
Night Vision and Electronic Sensors Directorate
10221 Burbeck Road
Fort Belvoir, VA 22060-5806

ABSTRACT: In the experiments reported here, trained military observers first searched for tracked targets in approximately 100 synthetically generated Paint The Night (PTN) images. Months later, a different group of trained military observers searched for tracked targets using imagery obtained with a particular advanced sensor operating in the 8 - 12 μ band. In both experiments, the maximum amount of time t_L observers were permitted to search an image was controlled and took on the values 3, 6, 9 and 12 seconds. The probability P(t) for the observers detecting the target as a function of time was measured for different values of t_L, for targets that varied continuously between easy and hard to find, and excellent agreement in search performance was observed when comparing PTN and imagery obtained in field trials utilizing an advanced thermal imager. The excellent agreement in search performance using PTN and field imagery, the ability to control target signature and backgrounds in PTN and the relatively low cost of search experiments done utilizing PTN imagery, make PTN an excellent tool for search model development.

1.0 Introduction and Overview.

Search model development at NVESD is always validated with field experiments and up to this point has also been largely developed from field tests. Field tests are expensive, it is difficult to control target signatures, it is difficult to get a variety of targets and one must use sensors that are available. These issues are neatly dealt with by using synthetic Paint The Night (PTN) imagery.

However, using PTN raises a question: Can observer search performance measured using PTN imagery be used to predict how observers will do using real thermal sensors searching for real targets in the field? The chief purpose of this report is to answer this question. The methodology utilized in this report involves comparing search performance for observers viewing PTN imagery with performance for observers using imagery collected in the field using an advanced 8-12 μ sensor. Future work will investigate the suitability of PTN for search model development with other sensors.

All experiments involving search will result in the calculation of different parameter values for detection time and probability of detection for the imagery presented. This makes a direct comparison of the raw data impossible. The method chosen here to accomplish this comparison utilizes interpolation functions to exactly fit experimental data obtained from PTN imagery. Then the interpolation function PTN results are compared against real sensor data. This method was chosen instead of the typical model development methodology because the goal was not to develop a model, rather to compare results from one experiment to that of another. The notation utilized in this report is therefore, by necessity, different from that utilized in the classical search literature and will now be described.

We are interested in how observer search performance depends on the maximum amount of time t_L observers are permitted to search an image. In the experiments reported here observers are permitted to search for a maximum of 3, 6, 9, 12 and 17 seconds.

P_∞ is one of the parameters usually used to describe search. It is the fraction of observers viewing an image who correctly declare a target given a large value for t_L. The parameter P_∞ measures how difficult it is for an observer to find a target in an image. A problem with this parameter is that the value of t_L is not specified in the definition of P_∞. Conceptually, to get an accurate value for P_∞, t_L is infinite. However, as t_L increases beyond 15 or 20 seconds observers get bored and performance suffers. To measure P_∞ , a compromise is made between the conceptual requirement for a large value of t_L and the practical

necessity of keeping t_L small enough not to bore the observers. Historically, P_∞ has been measured with a t_L of about 30 seconds. In the research reported here, the issue of what t_L to use in measuring P_∞ is sidestepped by introducing a different parameter ... Pe12.

Pe12 is the fraction of observers viewing an image who correctly declare a target given a maximum search time of twelve seconds. It measures how difficult it is for an observer to find a target in an image. The "e" in Pe12 stands for *experimental.* Pm12 denotes the *modeled* valued for Pe12.

The form of the search model for the experiments reported here is:

$$P(t) = a\left(1 - e^{-\frac{t - t_d}{\tau}}\right) \qquad t_d \le t < t_L \tag{1.1}$$

In equation (1.1) t_d is the average time observers taking the experiment took to move the mouse pointer to the vicinity of the target and click. P(t) is the fraction of observers who declare a target in time t. In this report, t is the time from when the image is first presented to the time an observer clicks the mouse. Search research workers frequently use t to denote what in this report is denoted by $t - t_d$. Observe that equation (1.1) only applies for t between t_d and t_L. The parameter τ is a time constant which determines how quickly the function P(t) reaches its asymptotic value. It may seem strange that the symbol "a" is used instead of the usual symbol P_∞. The reason for this is that in equation (1.1) "a" is a function that depends on Pe12 and t_L. Similarly, τ and t_d are functions of Pe12 and t_L.

$$\begin{aligned} a &= a(Pe12, t_L) \\ \tau &= \tau(Pe12, t_L) \\ t_d &= t_d(Pe12, t_L) \end{aligned} \tag{1.2}$$

Equation (1.1) with "a", τ and t_d given by equation (1.2) accurately describe observer search performance for observers viewing PTN imagery (see Figure 3.1). Equations (1.1) and (1.2) imply that P(t) depends on t with Pe12 and t_L as parameters. **What is shown in this report is that functions a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, tL) can be determined using PTN imagery and then these <u>same functions</u> can be used in equation (1.1) to accurately (see Figure 4.1) predict search performance when observers view imagery collected with the advanced sensor.** Future studies will investigate how robust this result is for different sensors, targets, backgrounds and observers.

Although the functions a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, t_L) are expressed algebraically, the expressions are long. The reason is that the functions a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, t_L) are bilinear interpolating functions. Space limitations preclude exhibiting these functions here. More concise expressions could have been written by function fitting using the method of least squares, but these would not represent PTN search data as accurately as the bilinear interpolation functions. It should be realized that for the most accurate comparison between PTN and sensors, different "a" and τ and t_d functions are used for 1st and 2nd gen sensors.

2.0 Experimental Procedure.

2.1 Image Generation.

The first part of the experiment was done using computer generated PTN imagery; the second part of the experiment was done on field imagery obtained with an advanced thermal sensor.

Briefly, PTN imagery is computer rendered imagery of real topography with computer generated rocks, trees, bushes and roads. PTN has the capability to place realistic vehicle models anywhere in the imagery. The apparent background and target temperatures as well as observer position are computer inputs. This enables a computer-generated image to be made at any time in the diurnal cycle from any observer position. PTN renders Long Wave Infrared targets and backgrounds and calculates a transmission loss by applying Beer's law to each pixel in the scene based on pixel range and an input attenuation per kilometer parameter. Sensor effects are then applied in a post rendering process, either real-time or in non-real-time. A more complete description of how PTN imagery is generated is given elsewhere [1 – 3].

A description of the particular imagery used in the experiments follows. Targetless, high resolution imagery, with trees, rocks, topography and roads typical of what is found at a U.S. Army site was generated for a particular region viewed from several positions on the perimeter of the region. Topography at the army site was sampled with an eight meter Cartesian grid before computer rendering. In each case the sensor was approximately five feet above ground level yet could see for a considerable distance because the sensor was located on the side of a hill. Nominally, one hundred high-resolution, pristine images were generated for several regions. Each PTN image had a field of view of 5.4 by 4 degrees, intermediate between that of a typical 1st and 2nd gen sensor.

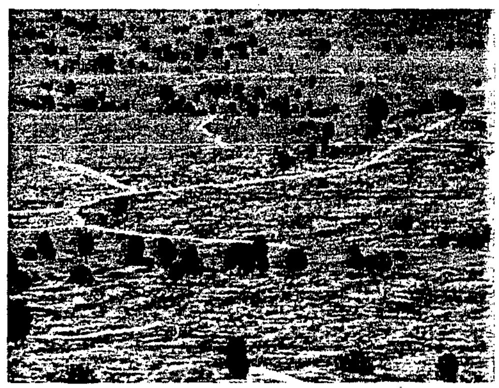


Figure 2.1. A representative advanced sensor image characterized by Pe12 ≅ 0.5 is shown on the left. Shown on the right is the corresponding second gen PTN image. Although a skilled observer has no trouble distinguishing an advanced sensor image from a PTN image, the PTN image is useful for preliminary search model development. The actual imagery seen by observers on monitors is better than that shown in this reproduction.

A single high-resolution target, selected from a set of targets was inserted into 68 of the images at realistic positions. The class of targets includes: M1A1, M60A3, HMMWV-Tow, BMP1 and an M113. Allowed target aspect angles include views at 0°, 30°, 150°, 270° and 330°. Aspect is defined as vehicle orientation relative to the sensor. Zero degrees correspond to looking at the target head on, ninety degrees corresponds to a direct view of the left side of the target and 270° corresponds to a direct view of the right side of the target.

The point spread function of 2nd gen imagery was generated by convolving: 1) the diffraction point spread function of a typical 2nd gen lens, 2) a Gaussian aberration blur function of a typical 2nd gen lens and 3) the spatial rectangle function associated with the finite size of a typical 2nd gen detector.

The PTN imagery was spatially sampled at a rate three times higher than the respective sensor resolution.

Each of the 100 images was convolved with point spread functions typical of a 2nd gen sensors as described above. No noise was added to the 2nd gen imagery. The 2nd gen imagery was DC coupled.

A goal in creating the PTN imagery was to have a range of easy and difficult targets so that the detection probability was uniform for 2nd gen sensors. This was accomplished by suitably adjusting target range, target signature strength and background signal strength.

Figure 2.1 compares a typical advanced sensor image with a simulated 2nd gen PTN image used in the perception experiments.

2.2 Power Point Presentation.

Observers viewed the PTN and advanced sensor imagery on individual monitors in groups of about seven over a period of several months. The number of observers who could take the perception test had to be seven or less because of the difficulty in fielding more than seven computers. To assure that observers received identical instructions in all sessions, instructions were given to the observers in a power point presentation.

A summary of these instructions follows. The observers were told a scene may or may not have a military vehicle in it. If a vehicle is present the observer's job is to find it; otherwise they were required to decide no vehicle is present as quickly and accurately as possible. Observers were shown a single image with five targets at five aspect angles displayed in a 5 x 5 matrix. They were then shown examples of representative targets seen with simulated 2nd gen imagery at different times of day, at different ranges and with different clutter.

Observers were also instructed in the mechanics of taking the test by the Power Point presentation. In particular, observers were shown the outline of a square that would move with the mouse and were told they would get credit for a correct detection if any part of the square overlapped any part of the target. This technique enabled the observers to quickly communicate perceived target location and is important in an experiment designed to measure detection time. Observers controlled when the next image was presented by pressing the space bar. When a new image appears, the square controlled by the cursor is automatically positioned at the center of the image. Observers were told this was a timed test and were encouraged to work as quickly as possible.

Definition of Detection. Observers were told if they saw something that they were uncertain about, but for which they would normally switch to a higher magnification, they should click on the object. If there are several items that satisfy this criterion, the observers were instructed to click on the one they thought was most likely to be the target. This defines detection for this test. This definition represents a change from the definition of detection [4] used previously at NVESD.

Definition of No Target Present. If the observer did not see a target and saw no object that they would examine more closely at higher magnification, then they were instructed to click the "No Target Present" button.

There was a practice session before taking the test and observers could practice as long as desired. They were also given a chance to ask the test administrator questions before taking the test.

2.3 The Experiment.

The observers were trained military personnel who were exceedingly good at finding targets. For some images, trained military observers could detect the target even when the authors of this paper had trouble seeing the target when cued to target location. The ability, displayed in this experiment, of the trained military observers to detect targets is awesome.

The experiment consisted of several sessions spaced nominally one month apart. This allowed us to determine that the experiment was well designed before having many observers take the test.

In the experiment, the observers could effectively make three choices: 1) if they thought they found the target, they could use a mouse to position a square box on the target and then click; 2) if they believed no target was present in the image they could click on a no target present box; 3) if they exceeded the time allowed to communicate a response, the current image would disappear. An observer caused the next image to be displayed by pressing a space bar.

The observers were given 3, 6, 9, 12 and 17 seconds to look at the imagery. They were also given as much time as they wanted to look at the imagery. Observers rarely needed more than 17 seconds to make a targeting decision. In those few cases where this occurred, it is believed this is due to an attention lapse. In a practical sense, 17 seconds corresponds to giving the observers as much time as they could productively use to find the target. The 100 images were shown to each observer in a random order without repeats i.e. like showing each card in shuffled deck. Each time the 100 images were shown, they were shown in random order. To assure the observers did not memorize the imagery, observers were shown the imagery first for 3 seconds, then for 6 seconds, then for 9, 12 and 17 seconds. In the final stage of the experiment observers could study the imagery as long as desired. Here we only report the data for 3, 6, 9 and 12 seconds. The reason: when observers were given more than 12 seconds they were fatigued. This compromised the credibility of data taken after 12 seconds.

Each observer took the experiment independently. A personnel computer was assigned to each observer for the duration of the experiment. The experiment took place in one dimly lit room.

PTN search sessions were done over the first three months. Advanced sensor sessions were done in the second three months.

2.4 Computer Hardware and Software.

The time to record a search choice depends on the quality of the mouse. For that reason, each observer was equipped with a recently purchased mouse and each mouse was tested to assure it was in excellent operating condition before it was used in the experiment.

Observers communicated with the computer by using the mouse to move a pointer and then clicking. When the pointer was on the image it assumed the shape of a square box with a 2 pixel wide line measuring thirty-six pixels on each side. This yields an area equal to approximately 1/300 the picture area. Observers indicated they found a target by positioning this square so that it overlapped any portion of the target. We were interested in how long it took the observers to find the target and did not want observers to spend an inordinate amount of time positioning the square on the target. For this reason, events where the observer clicked the mouse when any part of the square overlapped any part of the target count as correct. The observers were told this and they were urged to work as quickly and accurately as they could in a Power Point presentation before the test. When an observer believed no target was present, the mouse was moved to the side of the image where the pointer changed from a square to an arrow and the observer clicked a "No Target Present" button. If time ran out before the observer could communicate a decision, then a "Time Out" was recorded. The observer pressed the space bar to display the next image.

To get accurate detection times, mouse clicks are automatically recorded with the clock internal to the computer. Recently purchased personal computers and skillful programming assure these factors contribute a negligible error in the measured time it took an observer to make a decision.

3.0 Data Analysis.

To start with there are 100 PTN images and 100 advanced sensor images. In principal, each PTN image and each advanced sensor image could be individually analyzed but if that were done, with approximately 30 observers, the statistical error inherent in the search process dominates the results. For this reason, PTN and advanced sensor images were binned into four groups according to how difficult it is to acquire the target. The four PTN groups had Pe12 values of: 0.857, 0.560, 0.423 and 0.230. The four advanced sensor images had Pe12 values of: 0.86, 0.63, 0.47 and 0.32. Because the Pe12 values for PTN images are not identical to the Pe12 values for the advanced sensor images, P(t) values obtained using PTN imagery can not be directly compared with P(t) values obtained using advanced sensor imagery.

For a group of PTN images characterized by a particular Pe12 and t_L values, "a", τ and t_d values were determined by doing a least square fit to the experimentally determined P(t) vs t data. The Mathematica, Nonlinear Regress program that utilized the Levenberg-Marquardt algorithm, was used.

At this point we have tables of "a", τ and t_d as functions of Pe12 and t_L. One approach is to determine the equations (1.2) by assuming a functional form for these equations and then do a least squares fit. This approach was tried and was not completely satisfactory because it is difficult to know what functional form to assume. For the functional forms we assumed, the fit function matched the experimentally determined values of "a", τ and t_d with accuracies of about 10 %. However when these values where used in equation (1.1) the mathematical description of the experimental results was frequently off by 15 to 25 %. We were not satisfied with the accuracy with which the least squares approach described PTN experimental data and this caused us to consider a different approach.

Bilinear interpolation is the alternate approach. Abramowitz and Stegun [5] give the algorithm used. Bilinear interpolation reproduces the table of "a", τ and t_d as functions of Pe12 and t_L exactly and gives reasonable interpolations for these functions for Pe12 or t_L values that are not in the table. As illustrated in Figure 3.1, PTN search data was described mathematically by equations (1.1) and (1.2) with a high degree of accuracy. As illustrated in Figure 4.1, bilinear interpolation with parameters determined from PTN search experiments allowed accurate predictions of field search performance with an advanced sensor.

A graphical description of the functions a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, t_L) is exhibited in Appendix A. A description of the bilinear interpolation method and its virtues for the work described here is given in Appendix B.

4.0 Results.

Figure 4.1 shows a comparison of the mathematical model derived from PTN imagery and observer response to field imagery obtained with an advanced sensor. The agreement between the mathematical model and field imagery supports the view that PTN is an excellent tool for search model development.

5.0 Conclusions.

Equations (1.1) and (1.2) accurately represent search results when observers view PTN imagery and when observers view field imagery obtained with an advanced sensor (see Figures 3.1 and 4.1).

Measurements of observer response to PTN imagery allow accurate search predictions to be made for observers using an advanced thermal sensor (see Figure 4.1). Future studies will investigate how robust this result is for different sensors, targets, backgrounds and observers.

This investigation supports the view that PTN is a highly cost effective and useful tool for search model development. Of course a model developed from PTN imagery needs to be verified and validated with search experiments done on imagery collected in field tests.

Bilinear interpolation functions accurately summarized observer search response to PTN images (see Figure 3.1). This technique, rarely used in search modeling, was useful and is expected to gain wider use.

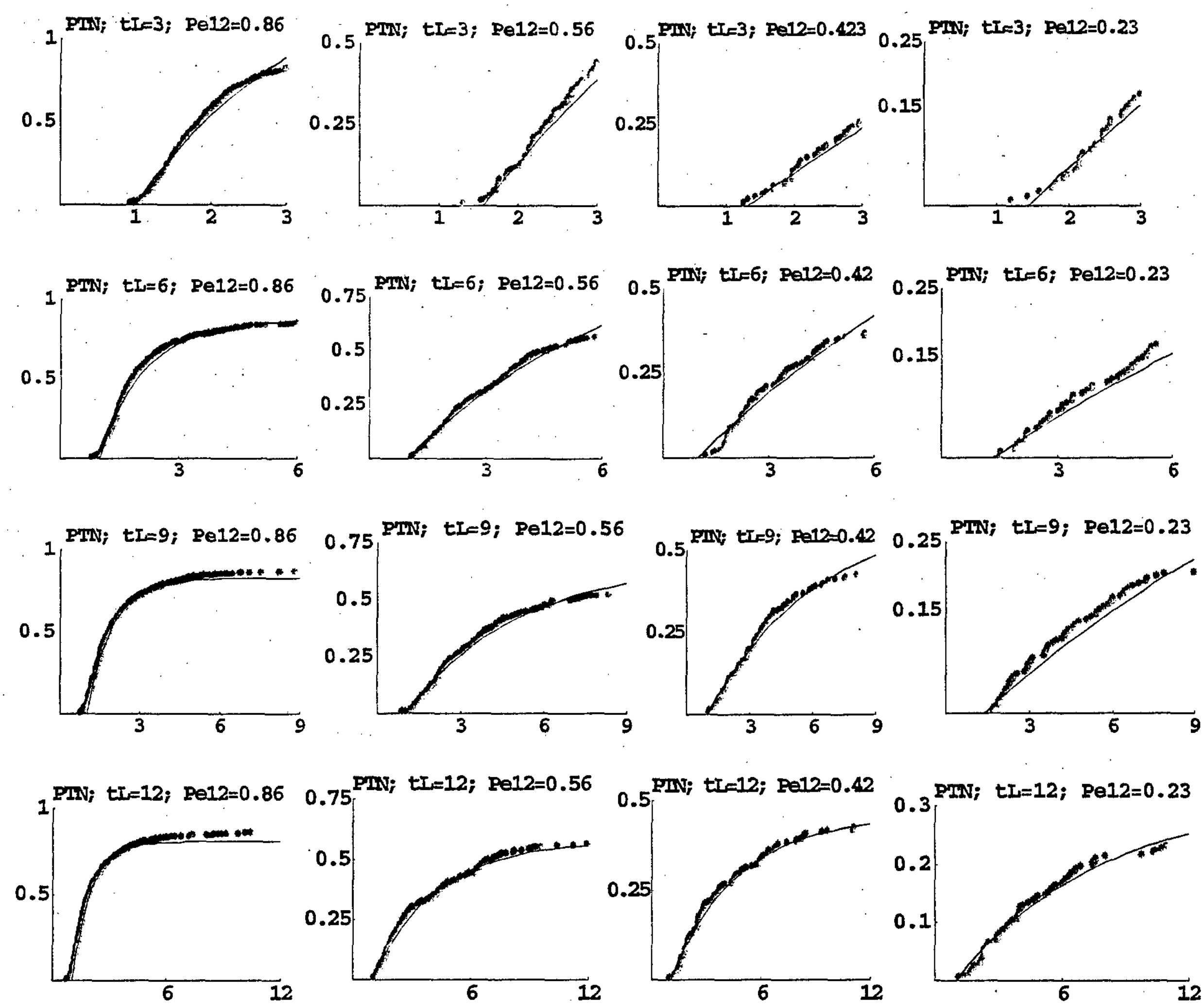


Figure 3.1. Shown above are graphs comparing a mathematical model, derived from observer response to PTN imagery, with the actual response of observers viewing PTN imagery. In each graph the horizontal axis corresponds to time and the vertical axis to probability of detection. The solid lines represent the results of the mathematical model and the dots represent observer response. As one goes from graph to graph, the probability of the observers detecting a target Pe12 changes in the horizontal direction; the maximum amount of time t_L observers are permitted to search an image changes in the vertical direction. The close agreement between the mathematical model and observer response demonstrates how well equations (1.1) and (1.2) describe observer response to PTN imagery.

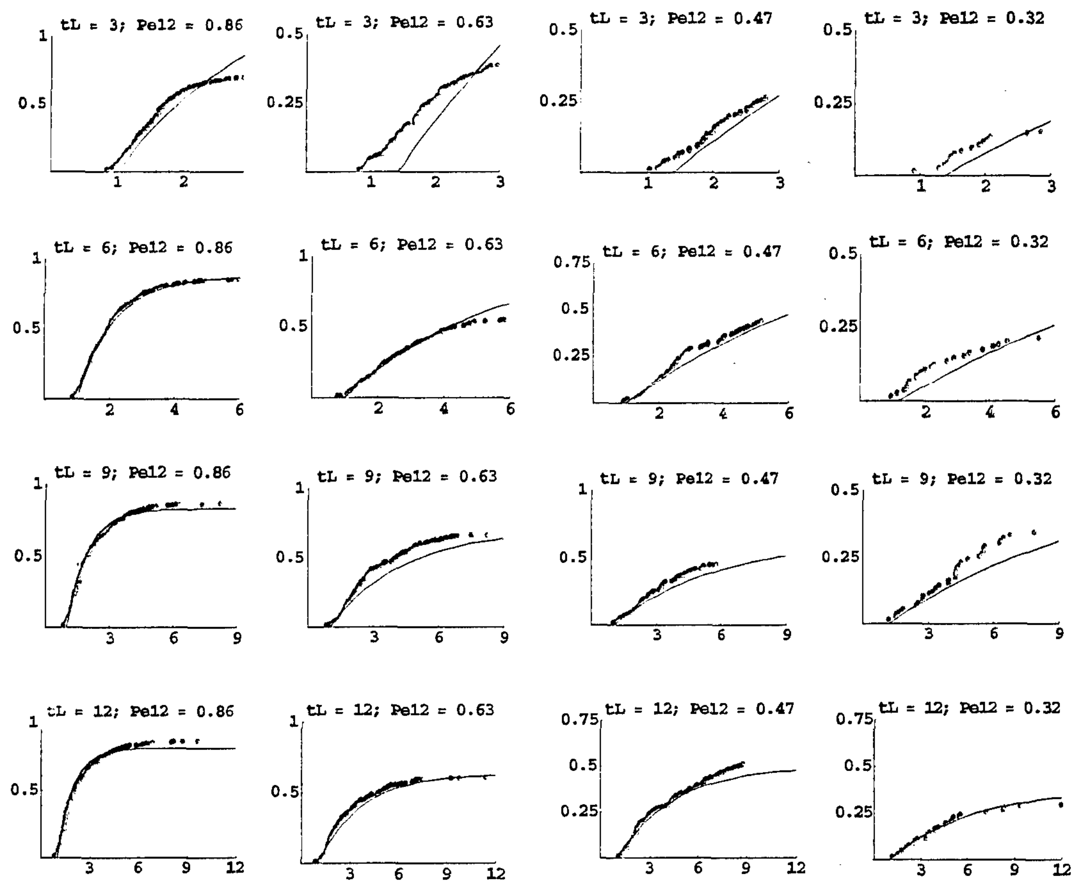


Figure 4.1. Shown are graphs of a mathematical model, derived from observer response to PTN imagery, with the response of observers using an advanced thermal sensor and real targets. In each graph the horizontal axis corresponds to time and the vertical axis to probability of detection. The solid lines represent the mathematical model and the dots represent observer response using an advanced imager. As one goes from graph to graph, the probability of the observers detecting a target Pe12 changes in the horizontal direction; the maximum amount of time t_L observers are permitted to search an image changes in the vertical direction. The agreement between the mathematical model and observer response using the advanced sensor is remarkable. The agreement between the mathematical model and field imagery supports the view that PTN is an excellent tool for search model development.

Appendix A. Graphical Representation of the Functions a(Pe12, t_L), τ(Pe12,t_L) and t_d(Pe12,t_L)

Using the bilinear interpolation method described in Appendix B, the functions a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, t_L) were determined from observer search response to PTN images using methodology described in section 3. Figure A.1 shows these functions graphically.

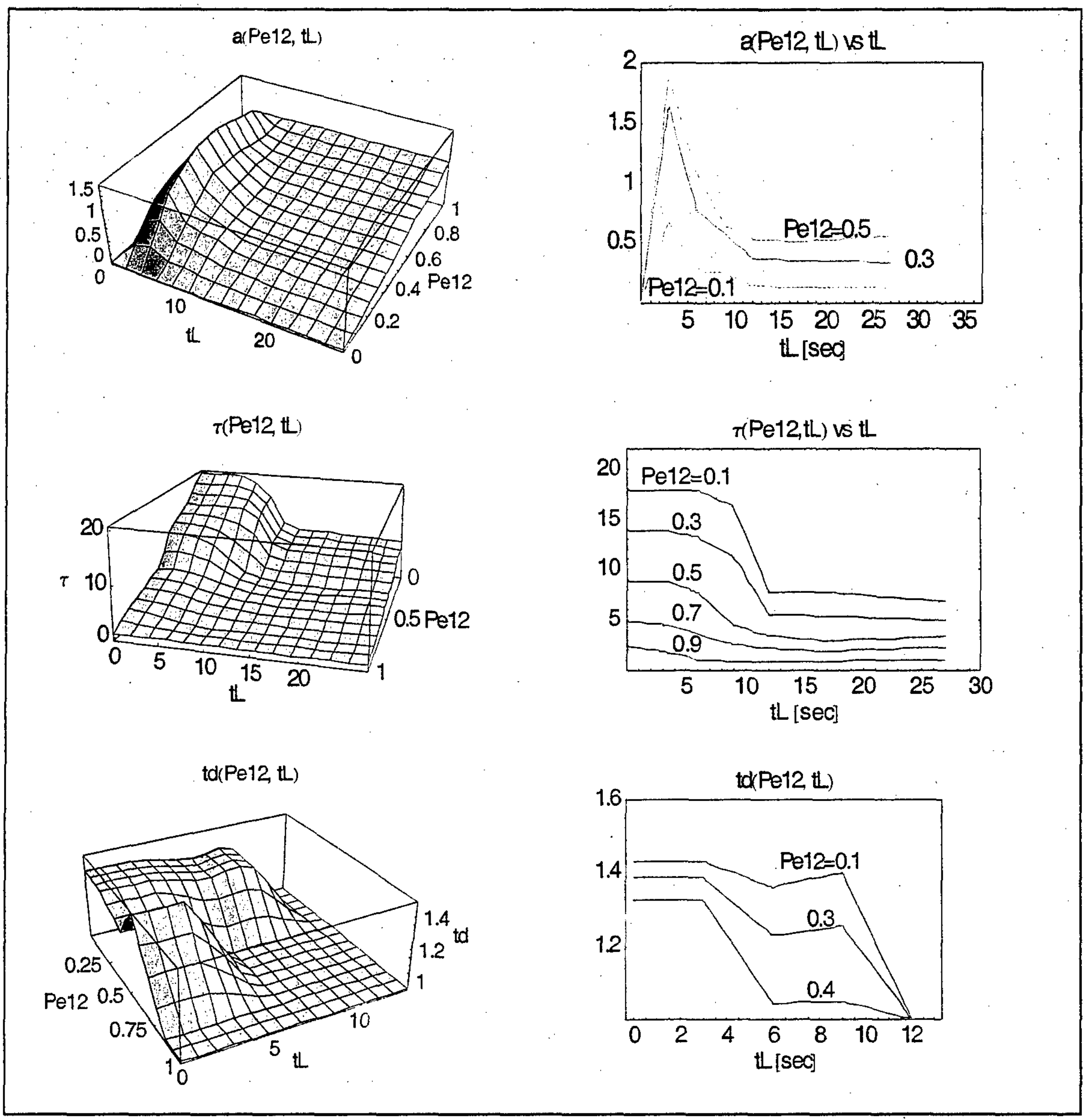


Figure A.1. Surfaces depicting a(Pe12, t_L), τ(Pe12, t_L) and t_d(Pe12, t_L) determined from observers viewing PTN images are shown on the left. Parametric curves for those surfaces are shown on the right.

Appendix B. Bilinear Interpolation and Its Virtues

The caption to Figure B.1 defines the problem which bilinear interpolation solves.

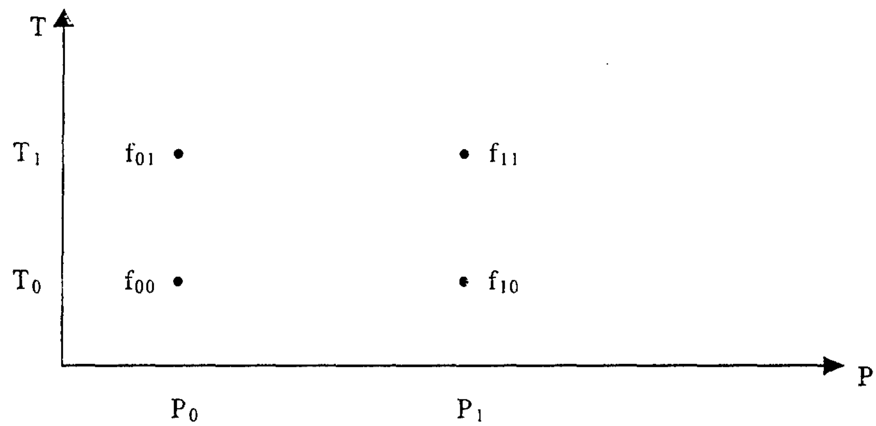


Figure B.1. In this figure P is an abbreviation for Pe12 and T is an abbreviation for t_L .The symbol f can be either "a", τ or t_d. With f_{00}, f_{10}, f_{01} and f_{11} known, the objective of bilinear interpolation is to estimate f when P is between P_0 and P_1 and T is between T_0 and T_1.

Define reduced variables p and t:

$$p \equiv \frac{P - P_0}{P_1 - P_0} \qquad t \equiv \frac{T - T_0}{T_1 - T_0} \tag{B.1}$$

Observe in equation (B.1) that as P varies between P_0 and P_1, p varies between zero and one. Similarly as T varies between T_0 and T_1, t varies between zero and one. The bilinear interpolation function f(P, T) is:

$$f(P,T) = (1-p)(1-t)\,f_{00} + (1-p)\,t\,f_{01} + p(1-t)\,f_{10} + p\,t\,f_{11} \tag{B.2}$$

The interpolating function (B.2) has desirable properties.

- f matches the data points exactly.
  - When p and t are zero, $f = f_{00}$
  - When p and t are one, $f = f_{11}$
  - When p equals one and t equals zero, $f = f_{10}$
  - When p equals zero and t equals one, $f = f_{01}$
- The center point averages the surrounding points.
  - When p equals ½ and t equals ½ , $f = ¼ (f_{00} + f_{10} + f_{01} + f_{11})$
- f is linear in p and in t.
  - This implies that if t is held constant, f interpolates linearly between data points with different p. Similarly if p is held constant, f interpolates linearly between data points with different t.
- The preceeding bullet implies the interpolating function produces a continuous surface.
  - This is important because a small change in p or t should not cause a discontinuous change in the probability for detecting the target.

REFERENCES

1. Lorenzo, M., Jacobs, E., Moulton, J.R., Jr., "Optimal Mapping of Radiometric Quantities in OpenGL," Proceedings of SPIE, April 1999.
2. Lorenzo, M., Moulton, J.R., Jr, "Fidelity Requirements for EO Sensor Simulations" Proceedings of the Spring 1997 Simulation Interoperability Workshop, March 1997
3. Lorenzo, M and Moulton, R. "Toward Standards for Interoperability and Reuse in IR Simulation", Proceedings of the Fall 1997 Simulation Interoperability Workshop, September 1997.
4. The standard definition of detection used at NVESD: A target is said to be detected when a trained military observer can determine with some confidence that the object is of military significance. See M. Friedman et al. "Standard Night Vision Thermal Modeling Parameters", Infrared Imaging Systems: Design, Analysis, Modeling, and Testing III, p 204, SPIE Proceedings, Vol 1689 (April 1992).
5. Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables, Editors M. Abramowitz, L. A Stegun, National Bureau of Standards Applied Mathematics Series 55 (1964). See equation 25.2.66.

Algorithm for the Iterative Design of Observer Field Tests

John G. Bennett
U.S. Army Tank-automotive and Armaments Command
Warren, MI 48317-9000

ABSTRACT

In a previous paper, I described a technique for designing an observer test in an iterative manner. In field tests to compare the observability of combat vehicles, the test designer must select the optimum number of observation opportunities in order to balance collecting enough data to draw valid conclusions against the high cost of supporting vehicles and personnel at a test site. The test designer can select the number of observations, N, so that a given experimental difference in detectability will be statistically significant at a given confidence level. Alternatively, the test designer can select N so that the probability of rejecting a given underlying difference in detectability is less than a given amount. The test designer, however, generally lacks key parameters for the efficient design of the test. Namely, the designer lacks the detection probabilities of the vehicles at each range. The standard deviation of the difference in detection probability depends upon the detection probability itself. Therefore, the test designer must select the number of observations for each range based upon the conservative assumption that the probabilities are near 50%, the probability for the maximum standard deviation. In the previous paper, an iterative technique of test design was described. In this technique, the test designer modifies the test matrix as the test progresses. Early test results yield estimates of the probability of detection for each vehicle at each range. Based on these estimates, the test designer reallocates the number of observations among the ranges, improving the efficiency of the test.

In this paper, I present an algorithm to implement this iterative technique. Initially, observation opportunities are equally distributed among the ranges of vehicle sites. But, as the test progresses, the observation opportunities are redistributed among the ranges according to estimates of detection probability. For example, the fraction of opportunities at the range near 50% probability is increased while the fraction near 10% probability is decreased. The algorithm handles the special cases that occur when the estimated probability is 1 or 0.

1. THE EXPERIMENTAL SITUATION

Figure 1 illustrates a typical test setup for a test of observability. Observers are stationed at a fixed site and attempt to detect a vehicle in their field of view. For each observation opportunity, the test personnel record the number of detections. In analyzing the data, the analyst groups the observations into range bins and compares the proportion of detections for each test vehicle.

For such a field test, the test designer must select the optimum number of observation opportunities at each range. The designer must balance collecting enough data to draw valid conclusions against the high cost of supporting vehicles and personnel at a test site.

2. DESIGNING THE EXPERIMENT

The test designer selects the number of observations in order to achieve two objectives. First, the number of observations must be large enough so that the probabilities of errors are below the desired level. And, second, the number of observations must be as small as possible to minimize testing time and cost.

Table I illustrates the types of errors that can occur in drawing conclusions from an observation test. The goal of an observation test is to judge whether the observed difference in detection probability, Pd, is unusual enough to reject the null hypothesis that Vehicle A and Vehicle B have the same Pd. If the analyst concludes that the null hypothesis cannot be rejected, then either the analyst has made a correct decision or a Type II error. On the other hand, a decision to reject the null hypothesis will be either correct or a Type I error. In terms of countermeasure effectiveness, a Type I error is an erroneous conclusion than an ineffective countermeasure is effective, while a Type II error miss judges an effective countermeasure.

In a previous paper [1], I discussed the number of observation opportunities required to control the probability of these errors. For the case of an average Pd of 0.5, Figure 2 plots the number of observations required to maintain the probability of Type I and Type II errors at less than 5%. For example, 85 observations of each vehicle are adequate for an observed 0.15 difference to be significant with less than 5% chance of a Type I error. On the other hand, for less than 5% chance of committing a Type II error when the underlying probabilities differ by 0.15, the requirement is 316 observation opportunities per vehicle.

3. AN ITERATIVE TECHNIQUE FOR TEST DESIGN

An iterative technique can reduce the number of observation opportunities required for a test. Observation data is described by the binary distribution. But the binary distribution has the property that the standard deviation depends upon the probability. If p is the probability and N is the number of trials, then the standard deviation of the number of successes, σ, is given by

$$\sigma_{\text{detections}} = \sqrt{Np(1-p)} \tag{1}$$

And the standard deviation of the proportion of successes is given by

$$\sigma_{\text{proportion}} = \frac{\sigma_{\text{detections}}}{N} = \sqrt{\frac{p(1-p)}{N}} \tag{2}$$

Figure 3 plots Equation 2 for an N of 100, 200 and 300. Figure 4 plots the same equation normalized to its maximum at p = 0.5.

Figure 5 further shows how the number of trials changes with probability if the standard deviation is held fixed. For example, in comparison to a probability of 0.5, only 60% as many trials are need at a probability of 0.1.

To use this property of the binary distribution, the test designer modifies the test matrix as the test progresses. Early test results yield estimates of the probability of detection for each vehicle at each range. Based on these estimates. the test designer reallocates the number of observations among the ranges, improving the efficiency of the test design. Figure 5 illustrates a sample of the improvement in efficiency that this technique can achieve. Initially the test designer would have selected the number of observations at each test range by assuming the worst-case value of Pd = 0.5. But as the test progresses, the designer would use estimates of Pd from the early measures to redesign the test. For example, at range 20, the designer would reduce the number of observations from 314 to 251. Overall in this sample test, the test designer would reduce the number of observations from 1570 to 1193, a reduction of 24%.

4. IMPLEMENTATION OF THE ITERATIVE TECHNIQUE FOR TEST DESIGN

A. Algorithm for Fixed Number of Observation Opportunities

The first step is selection of the number of observation opportunities for each range based on the conservative assumption of a Pd of 50% at each range. From figures x and y, the designer selects No to reduce the probabilities of Type I and II errors to acceptable levels.

Next, the designer calculates estimates of P_d^i, the average probability of detection at range R_i, based on the results of the first day of observations, as follows:

$$P_d^i = \frac{D_A + D_B}{N_A + N_B} \qquad (3)$$

where D_A and D_B are the numbers of detections of Vehicles A and B and N_A and N_B are the number of observation opportunities for Vehicles A and B.

To account for special cases when there are no detections or 100% detections, Equation 3 should be modified to prevent later setting the number of observation opportunities to zero. The following form covers these special cases,

$$\begin{gathered}\text{If } (D_A+D_B)/(N_A+N_B) < 0.1\text{, then } P_d^i = 0.1\\ \text{If } (D_A+D_B)/(N_A+N_B) > 0.9\text{, then } P_d^i = 0.9\\ \text{Otherwise, } P_d^i = \frac{D_A + D_B}{N_A + N_B}\end{gathered} \qquad (4)$$

Next, beginning with Equation 3, the designer calculates N_i, the new number of observation opportunities for range R_i:

$$\sigma = \sqrt{\frac{P_d^i(1-P_d^i)}{N_i}} \qquad (5)$$

Note that σ is the same for all ranges. Solving for N_i yields,

$$N_i = \frac{P_d^i\left(1-P_d^i\right)}{\sigma^2} \tag{6}$$

And, keeping the total number of observation opportunities constant at N_T, then

$$N_T = \sum_{i=1}^{R} N_i = \frac{1}{\sigma^2}\sum_{i=1}^{R} P_d^i\left(1-P_d^i\right) \tag{7}$$

Solving for σ^2, gives

$$\sigma^2 = \frac{1}{N_T}\sum_{i=1}^{R} P_d^i\left(1-P_d^i\right) \tag{8}$$

And

$$N_i = N_T \frac{P_d^i\left(1-P_d^i\right)}{\sum_{j=1}^{R} P_d^j\left(1-P_d^j\right)} \tag{9}$$

Now, the test continues with the modified number of opportunities for observation at each range. At the end of each succeeding day, the designer can recalculate the N_i's in the same way to refine the design.

B. Algorithm for Minimal Number of Observation Opportunities

As an alternative to keeping the total number of opportunities for observation fixed, the designer can decrease the number of opportunities for P_d's that differ from 50%. With this criterion, the new number of observations can be calculated by using Equation 2 to keep the standard deviations constant,

$$\sqrt{\frac{P_d^i\left(1-P_d^i\right)}{N_i^{new}}} = \sqrt{\frac{0.5(1-0.5)}{N_i^{initial}}} \tag{10}$$

Solving for N_i^{new} gives,

$$N_i^{new} = \frac{N_i^{initial}}{0.25} P_d^i\left(1-P_d^i\right) \tag{11}$$

Again, at the end of each succeeding day, the designer can recalculate the N_i's to refine the design.

5. CONCLUSION

In this paper, I have presented an iterative technique of using early test results in an observation test to improve the overall efficiency of the test design.

6. REFERENCE

1. J. G. Bennett, " Proceedings of the 2001 Ground Target Modeling & Validation Conference, Houghton, MI, August 2001.

Possible Outcomes of Hypothesis Testing

	Is Vehicle A Less Detectable Than Vehicle B?	
Decision:	**No**	**Yes**
Accept Null Hypothesis	Correct	Type II Error
Accept Alternative Hypothesis	Type I Error	Correct

Table 1. Definitions of Type I and Type II errors.

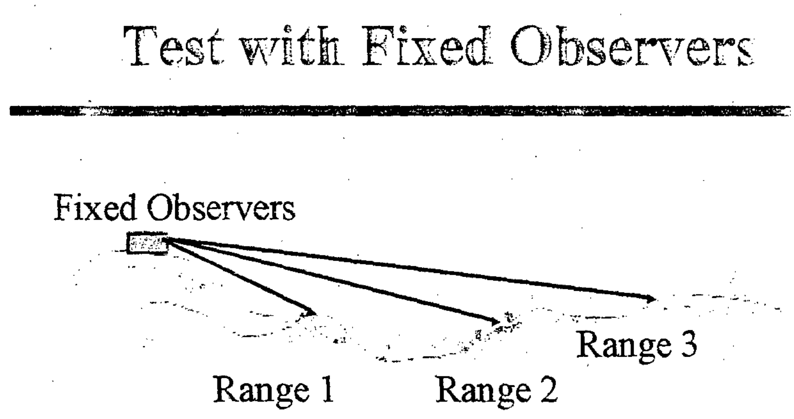


Figure 1. Experimental setup for an observer test.

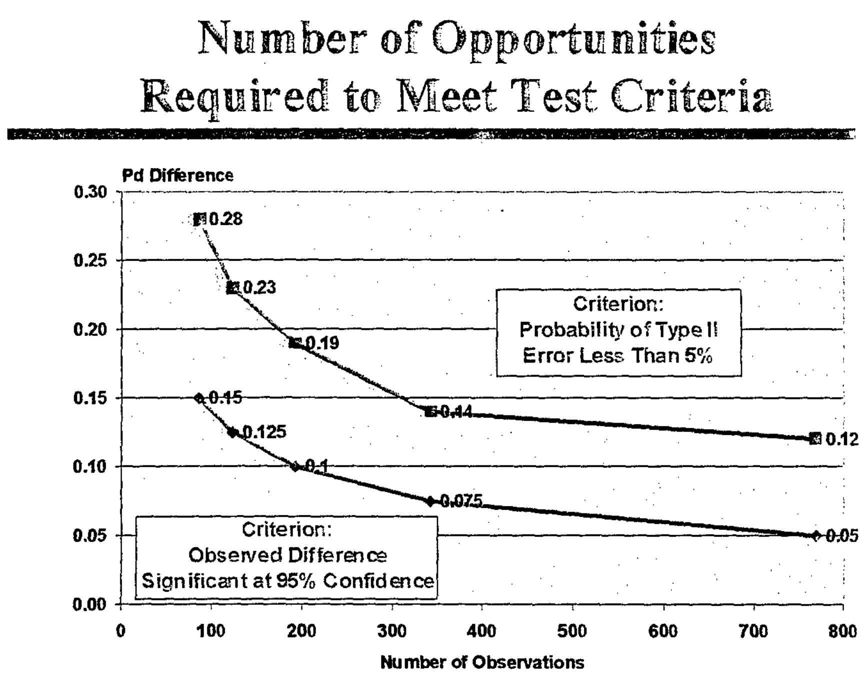


Figure 2. Number of observation opportunities required to meet test criteria.

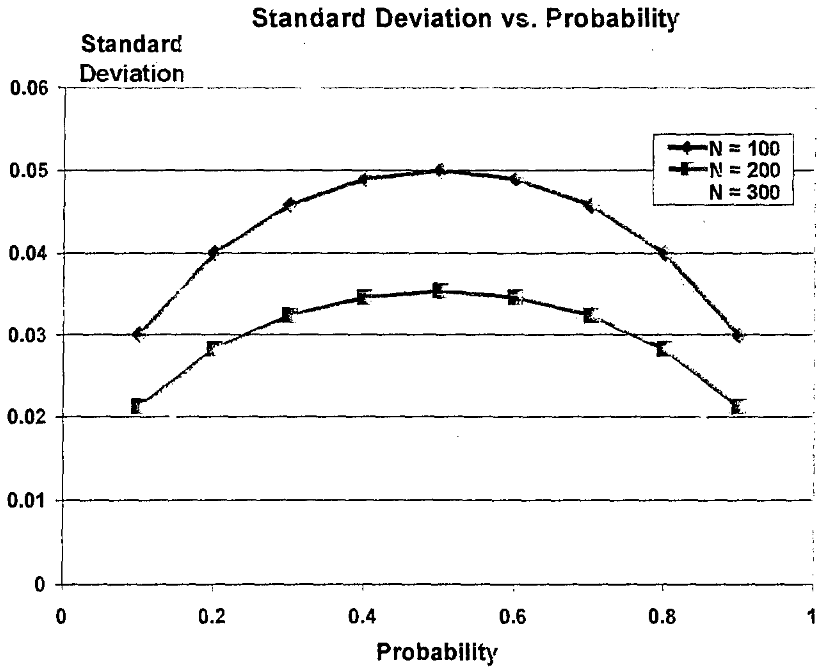


Figure 3. Standard deviation of a binary distribution for 100, 200 and 300 trials.

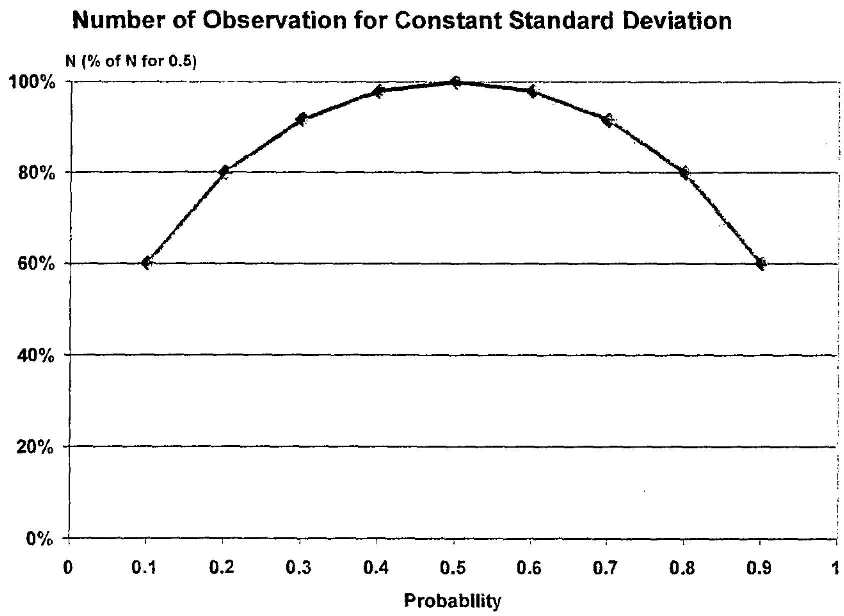


Figure 4. Standard deviation of a binary distribution normalized to probability of 0.5.

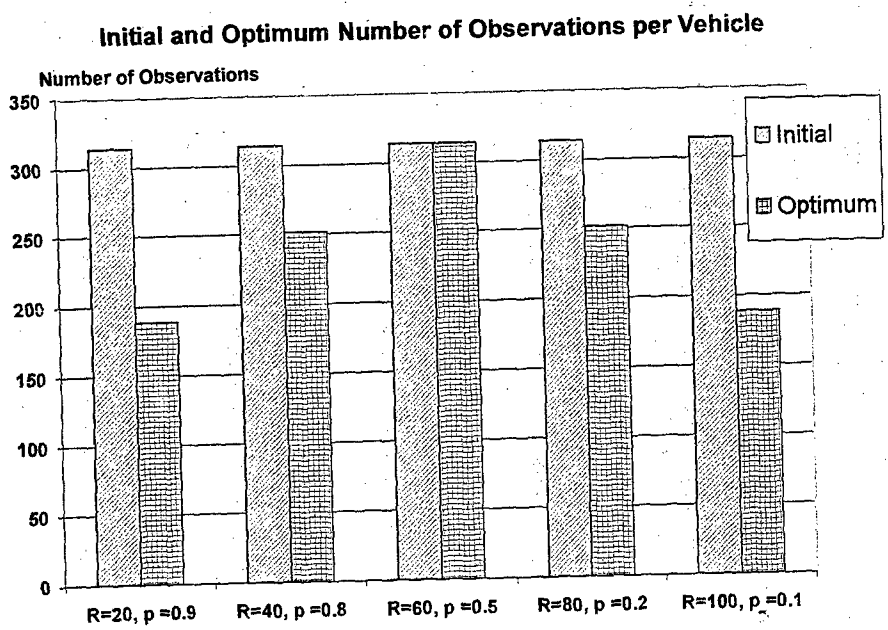


Figure 5. An example of iterative design of an observer test.

Validation of a New Analysis Program for Law of Comparative Judgment Data

by
James J. Crile, US Army TARDEC
and
James R. McManamey, US Army NVESD

ABSTRACT

The psychophysical Law of Comparative Judgment (LCJ) has been shown to have substantial value in assessing the signature of camouflaged or signature managed (C/SM) military assets when there is a need for a psychophysical measure of effectiveness. However, statistical tools for analyzing LCJ data are not widely available, have until now only had command-line interfaces, and require extensive interaction with the analyst. This prompted the Army to produce a new computer program that combines the functionality of several older command-line programs and has a more convenient windows interface. Exercising normal caution with new computer code, it is necessary to verify the new program to ensure that it correctly performs the desired tasks. This process requires a number of test cases: examples of each different class of results to be expected. Additionally, some of the changes in the program are expected to affect the precision of values for LCJ scales. Thus, since the values obtained with the new program are expected to be different than those obtained with the old one, it is necessary to verify that the results are equivalent. This paper describes the nature of the changes, the validation process, the nature of the validation data, and the results of the validation effort. The conclusion states that the new program does give results that are very similar to those obtained with the old programs.

BACKGROUND

The Law of Comparative Judgment (LCJ) is a method of psychological scaling that was introduced by Thurstone in 1927[1] and refined thereafter. It was summarized in 1958 by Torgerson[2], a staff member of the Lincoln Laboratory at Massachusetts Institute of Technology. This was a seminal work, explaining, among other things, both the theory and methodology of calculating an LCJ scale. It was of sufficient importance that it was reprinted in 1985, although it is now out of print[3]. Without a computer of substantial capacity, calculating an LCJ scale is time consuming and error prone, but the rapid growth of computer speed and memory capacity in the 1970's removed the computational impediments to its use, with the exception of the lack of suitable software. In the period from 1993 to 1997 that impediment was removed when Copeland, et al., adopted the LCJ for their X-based Perceptual Experiment Testbed (XPET)[4]. Since that time, the LCJ has enjoyed a surge of popularity in military evaluation of target signatures[5, 6, 7].

An LCJ Scale as a Model

An LCJ scale is actually a statistical model of the raw data that it represents. The data are viewed as being a function of a single variable that describes the intensity of a psychophysical quantity. This quantity may be any perceived quality, tangible or intangible, such as loudness, beauty, or distinctness. Even if the quality has a tangible, physical component, such as loudness, the LCJ is concerned with the psychophysical perception of that quality. Since there is no way to directly measure the intensity of this internal, psychophysical quality, the subjects are simply asked to decide which of two stimuli has more of the quality being rated. The stimuli are presented either at the same time in close physical proximity to each other or in the same place in close temporal proximity to each other. By using appropriate statistical methodologies, one may process the data from a suitable assessment session and arrive at the scale values. The scale values are then predictors of the raw response data in much the same way that a regression line is a predictor of data that is a function of two variables, one dependent and one independent. For example, suppose that I fill the gas tank of my car. Then, after driving 344.1 miles, it needs 14.8 gallons of gasoline to fill it back up. I repeat this process 3 more times as shown in Table 1. The result is shown

in the graph in Figure 1. The graph also shows a trend line found by linear regression. It is a model of gasoline consumption in my car. It shows that, on the average, I am able to drive about 23.3 miles for each gallon of gasoline that I use, plus about 8.0 additional miles per tank-full. If I assume a linear relationship between miles driven and gallons used, this is the best model for fuel consumption of my car based on this information.

Table 1 – Fuel consumed for various distances driven in a fictitious automobile.

gallons	miles
14.8	344.1
15.3	372.6
17.1	400.6
16.5	395.8

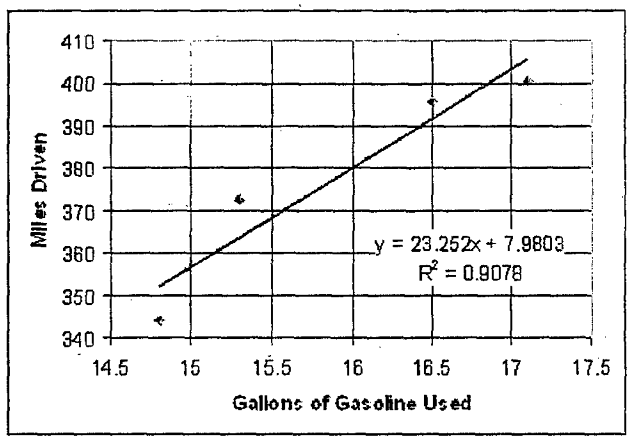


Figure 1 – Four observations of fuel consumption and associated miles driven (Table 1) along with the regression line that best describes an assumed linear relationship between the two quantities.

In the same way, the LCJ develops the linear model that best explains the responses to a set of stimuli when they are rated on some trait. The difference is that the input for calculating an LCJ scale is a two-dimensional array of responses (rating stimuli for relative values of a trait) and the output is the position of each stimulus on a linear scale for that trait. Just as with the scattered points in the gasoline consumption example, where the linear relationship is not at once easily identified, so the positions of the stimuli on a linear scale for the trait are also not easily determined. For example, suppose that 6 camouflaged targets are shown to 38 observers. They are shown in pairs. Each time a pair is shown, the observers pick the one that is the best camouflaged (hardest to see). Table 2 indicates the number of observers who pick each of the targets as best. We can easily see that A was chosen as better than E in 36 out of the 38 times when they were compared and better than F in all 38 of the cases where A and F were compared. Similarly, B was chosen as better than E in 34 out of 38 instances and better than F in 32 out of 38 instances. Thus, we know that A and B are better than E and F.

Table 2 -- Frequency with which each of 6 targets was chosen as better camouflaged when compared to each of the remaining targets.

		Chosen "Worse"					
		A	B	C	D	E	F
Chosen "Better"	A	0	23	28	33	36	38
	B	15	0	18	27	34	32
	C	10	20	0	26	33	28
	D	5	11	12	0	28	25
	E	2	4	5	10	0	17
	F	0	6	10	13	21	0

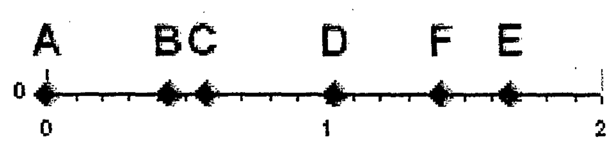


Figure 2 – Effectiveness of camouflage on 6 targets as judged by 38 observers (see Table 2). Zero is arbitrarily assigned to the most effective and higher values to less effective camouflage.

Although one can readily judge from these results that A and B are better than E and F, fine distinctions are difficult and developing a scale to know how much better any one is than any other is even more difficult. However, the statistical prescription associated with the LCJ develops a best-fit linear scale (shown in Figure 2). Using this scale, we can readily see

that treatment A is much better than treatment B. but B and C are very close to each other in effectiveness. The details of the statistical process that yields this scale are beyond the scope of this paper. For details, the reader is referred to the sources previously cited.

Need for new software

While the software included in XPET provides the means to accurately collect, combine, and process data from an LCJ evaluation session and produce a scale, there are still numerous steps. Additionally, the investigator is required to type command-line instructions in order to achieve this. Thus, to say that XPET is quick requires that one consider the older alternative of pencil, paper, statistical tables, and desk calculator. Also, practical limitations associated with field evaluations often result in an incomplete scale. Incomplete scales represent the most troubling limitation associated with the XPET software. We shall elucidate.

An incomplete matrix results whenever a cell that is not on the main diagonal contains a zero. This can occur for any of a number of reasons. A full discussion of the issue is beyond the scope of this paper, but if there are too many such cells, it also results in an incomplete scale. This means that it is not possible to produce a single scale and position all of the stimuli on it. When the scale is incomplete, one must take care that a division by zero does not occur while executing the scale development algorithm. The original XPET software was written in C and compiled for Silicon Graphics IRIS platforms. Due to some combination of default compiler options, library routines, and system features, when a division by zero is attempted, the program does not crash as one would ordinarily expect. Instead the result is flagged as NAN (Not A Number) and calculation proceeds. Subsequent calculations in which such a value is an operand, similarly result in a value that is similarly flagged. In the end, the results are printed with some of the results labeled as NAN. Copeland took advantage of this feature and did not trap division by zero errors, but instead allowed the defaults to handle them. While this was expedient and sufficient for the cases that Copeland encountered, it did not prove to be satisfactory in many instances encountered later. It has been especially problematic when the program that generates the scale is recompiled for systems that do not implement this error handling methodology.

Between the problems associated with calculation of incomplete scales and the desire to improve the user interface, it became desirable to rewrite some of the LCJ software. The authors of this paper served as a software development team to produce a new analysis program, based on certain components of XPET. In particular, the new software was to accept individual observer response files and files indicating stimulus presentation order as input and produce an LCJ scale as output. A "sum matrix" that summarizes all of the responses is a necessary intermediate product of the analysis. It was determined that such a sum matrix should be an available secondary output and an alternate input. In the course of developing the new software, a continuous function approximation was substituted for the result of trapezoid-rule numerical integration employed in XPET.

The Need to Verify the New Software

The new program functionally combines the following XPET utilities:

    xpet_pairs_result_2_matrix
    xpet_add_matrices
    xpet_pairs_lcj

Response matrices for individual observers (output of xpet_pairs_result_2_matrix) are optional. The sum matrix (output of xpet_add_matrices) may be retained along with the LCJ scale (output of xpet_pairs_lcj). Together, these last two outputs form the most compact and useful summary of an LCJ assessment, the sum matrix preserving an exact summary of the composite data and the LCJ scale being the best linear approximation. However, the creation of new software dictates that careful verification must be performed to ensure that equivalent results are obtained. In this case, equivalent results mean:

1. Given the same individual response files, the same sum matrix will be obtained, and
2. Given the same sum matrix, an equivalent LCJ scale will be obtained.

However, improvements incorporated in the new program make it likely that small differences will occur in the resulting LCJ scales. Furthermore, to say that one LCJ scale is equivalent to another does not require that they contain the same values. An LCJ scale is an interval scale[2,3]. As such, addition of a constant or multiplication by a constant does not substantially alter the scale. Thus, to have a high correlation between two scales is sufficient. The balance of this paper is concerned with the

procedures and findings whereby it was established that the new program is equivalent to the three XPET utilities it was designed to replace, or more precisely, that when the old and new programs process the same sum matrix, an equivalent scale is obtained.

VERIFICATION PLAN

Therefore, in order to verify that the new program operates as it should, we must first identify distinct modes of execution of the algorithms themselves. Once this has been accomplished, the next step is to obtain realistic decision matrices that cause such paths of the algorithm to be executed. We must then generate LCJ scales using both the new and the old programs. At this stage, the scale obtained using the new (Crile) program will be compared to the scale obtained using the old (Copeland) program. A high correlation will ensure that the programs are giving equivalent scales. Pearson's correlation coefficient will be used to indicate the degree of similarity between the scales. Then, each scale will become input to another program that generates a predicted response matrix. For each response matrix, we will compare the predicted matrix to the true decision matrix. This step will indicate which of the two scales is the best model of the original data. The objective of this comparison is to determine whether either program consistently produces a scale that is a superior model of the raw responses. The mean squared error will be used to indicate how similar two matrices are. The verification plan is illustrated in Figure 3 below.

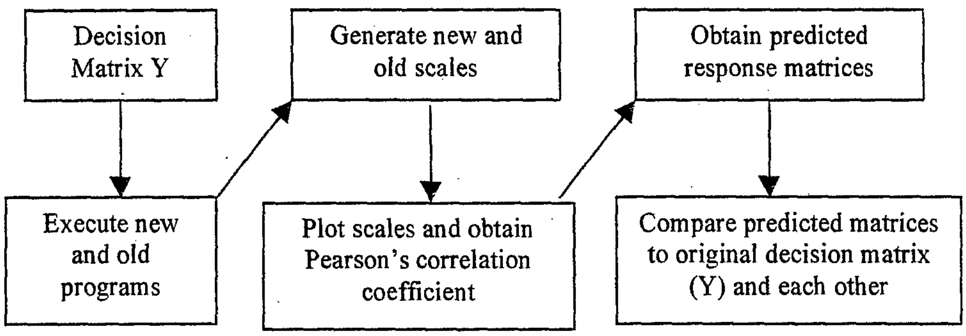


Figure 3 – The LCJ Program Validation Plan

In order to evaluate the successfulness of this verification, we will first need to establish some acceptance criteria. In general, we would say that a correlation coefficient of 0.95 between the two LCJ scales would be sufficient, given the inherent inconsistency that observers exhibit[8]. However, we would prefer to have a correlation coefficient of 0.99 or better, since purely mathematical operations account for the observed differences between the scales produced by the two programs and the observer data (program inputs) are precisely the same. For comparing the predictions of the two scales, there need be no upper limit on the sum squared error, since that is a function of the data. What we require is that this error be roughly equivalent in the two cases and that the error not be consistently higher with the new program. Using this method, we shall see if the acceptance criteria have been met. In doing so, we will determine whether or not the verification is successful. Having planned the verification process, let us now proceed to implement it.

Step 1: Identifying Distinct Cases

In this section, we will explain the four types of scales that can result from the LCJ program. At the same time, we will explain the basics of the algorithm.

The first thing we do is input a frequency matrix like Table 2 above into the LCJ program. This matrix is converted to a probability matrix (P-matrix) with the following equation for each element in the matrix:

$$P(i,j)=M(i,j)/(M(i,j)+M(j,i)), \tag{1}$$

where P(i,j) is an element of the P-matrix and where M(i,j) and M(j,i) are elements of the frequency matrix. For example, M(5,6) is 17 in Table 2, and M(6,5) is 21. Thus, P(5,6) is 0.477 and P(6,5) is 0.553.

Now, a P-matrix can be either complete or incomplete. The P-matrix is complete if it contains zeros only on the diagonal and all other entries are non-zero. If the P-matrix is incomplete. then it contains non-diagonal zero elements. At program execution, a complete P-matrix will always result in a normal scale. The frequency matrix type M1 (see Table 3) causes such an execution path. However, we will see that there are also other possibilities.

The next step in the program is to create an X-matrix. which is based on the cumulative normal standard deviate for each non-diagonal element in the P-matrix. P-matrix elements on the diagonal are ignored. Any other elements of zero or one will cause undefined values to occur in the X-matrix, since the cumulative normal standard deviate is defined only in the open interval from zero to 1 (excludes endpoints). If the X-matrix contains undefined values, then certain operations are performed on the columns to create a matrix of differences (D-matrix). The D-matrix can also have undefined values, and sometimes entire columns of it are undefined. The exact process behind this is beyond the scope of this paper, but let us be content with this explanation and continue.

From this point, identifying the other three cases is rather simple. If the P-matrix is incomplete but the D-matrix contains no undefined columns, then the matrix was of type M2. The scale resulting from such a matrix is equivalent to the scale of M1 except that it is arbitrarily chosen to begin with zero. If the P-matrix is incomplete and the D-matrix does contain undefined columns. then we end up with two or more unrelated scales called disjoint scales. Matrix type M3 contains two of these scales, and matrix type M4 contains multiple scales also of this sort. The four types are summarized in Table 3.

Table 3-- The Four Types of LCJ Results

Matrix Type	P-matrix	Scale
M1	complete	normal
M2	incomplete	normal
M3	incomplete	two disjoint scales
M4	incomplete	multiple disjoint scales

Step 2: Obtaining Exemplar Data

In Figures 4 - 7 below, we can see the four input matrices that were chosen for the validation. These matrices will yield P-matrices and scales as shown in Table 3. For example, M1 will result in a complete P-matrix and normal scale.

Step 3: Generated Scales

In this section, we will proceed to show the LCJ scale(s) that result from the four above matrices with both programs. We treat the old and new programs as "black boxes," showing the outputs without examining the method of computation. Some points will be made along the way, but for a more in-depth discussion, see **Old Model Versus New Model** (below).

In the M1 case (Figure 8), we can see that stimulus 6 received the highest rating and stimulus 4 received the lowest rating. Moreover, the lowest value is negative. There is no special significance to a negative value since these are interval scales (no natural zero) and adding any constant to every value leaves the scale, in essence, unchanged.

In the M2 case (Figure 9), we first notice that the lowest stimulus in each case is ranked zero—it is not negative as with M1. As stated above, when the P-matrix is incomplete, the first value in the scale is arbitrarily assigned zero.

We can see in Figure 10 (the M3 case) that we have two disjoint scales. This occurs because some of the stimuli are sufficiently different than the others so that the gap between two adjacent stimuli is essentially infinite (i.e. beyond measure by LCJ methods). This infinite gap can be observed in M3 (Figure 6) by the responses in the upper right and lower left quarters. We have recognized this by placing a blank line between the scales (Figure 10). The first scale contains LCJ values for 5 of the 9 stimuli. Here they are actually stimuli 1 through 5, but in general this need not be the case. The second scale

M1

0	218	230	82	166	211	197	172	220
70	0	124	37	70	179	163	137	190
58	164	0	29	85	184	168	143	195
206	251	259	0	209	160	143	117	171
122	218	203	79	0	224	211	186	232
77	109	104	128	64	0	121	104	57
91	125	120	145	77	167	0	119	59
116	151	145	171	100	184	169	0	88
68	98	93	117	56	231	229	200	0

Figure 4 – Response matrix that yields a complete P-matrix and a normal scale.

M2

0	14	14	18	14	16	12	17	20	12	13	18	19	22	17	19	14	20	22	16
8	0	12	16	12	14	10	15	19	10	11	16	17	21	15	18	12	19	22	13
8	10	0	15	11	13	9	14	19	9	10	15	17	21	14	17	11	19	22	13
4	6	7	0	7	9	5	10	16	5	6	11	13	19	10	14	7	16	21	9
8	10	11	15	0	13	9	14	19	9	10	15	17	21	14	17	11	19	22	13
6	8	9	13	9	0	7	12	18	7	8	13	15	20	13	16	9	17	21	11
10	12	13	17	13	15	0	16	20	11	12	17	18	21	16	19	13	20	22	15
5	7	8	12	8	10	6	0	17	6	7	12	14	20	11	14	8	16	21	9
2	3	3	6	3	4	2	5	0	2	2	6	8	16	5	8	3	11	19	4
10	12	13	17	13	15	11	16	20	0	12	17	18	21	16	19	13	20	22	15
9	11	12	16	12	14	10	15	20	10	0	16	18	21	15	18	12	19	22	14
4	6	7	11	7	9	5	10	16	5	6	0	13	19	10	13	7	15	21	8
3	5	5	9	5	7	4	8	14	4	4	9	0	18	8	11	5	14	20	7
0	1	1	3	1	2	1	2	6	1	1	3	4	0	2	4	1	6	15	2
5	7	8	12	8	9	6	11	17	6	7	12	14	20	0	14	8	16	21	9
3	4	5	8	5	6	3	8	14	3	4	9	11	18	8	0	5	13	20	6
8	10	11	15	11	13	9	14	19	9	10	15	17	21	14	17	0	19	22	13
2	3	3	6	3	5	2	6	11	2	3	7	8	16	6	9	3	0	19	4
0	0	0	1	0	1	0	1	3	0	0	1	2	7	1	2	0	3	0	0
6	9	9	13	9	11	7	13	18	7	8	14	15	20	13	16	9	18	22	0

Figure 5 – Response matrix that yields an incomplete P-matrix and a normal scale

M3

0	218	230	82	166	288	288	288	288
70	0	124	37	70	288	288	288	288
58	164	0	29	85	288	288	288	288
206	251	259	0	209	288	288	288	288
122	218	203	79	0	288	288	288	288
0	0	0	0	0	0	121	104	57
0	0	0	0	0	167	0	119	59
0	0	0	0	0	184	169	0	88
0	0	0	0	0	231	229	200	0

Figure 6 – Response matrix that yields an incomplete P-matrix and 2 disjoint scales.

M4

0	225	221	90	161	0	0	0	0
63	0	138	30	76	0	288	288	288
67	150	0	32	81	0	288	288	288
198	258	256	0	212	0	288	288	288
127	212	207	76	0	0	288	288	288
0	0	0	0	0	0	127	102	53
0	0	0	0	0	161	0	118	65
0	0	0	0	0	186	170	0	86
0	0	0	0	0	235	223	202	0

Figure 7 – Response matrix that yields an incomplete P-matrix and multiple disjoint scales.

contains LCJ values for the remaining 4 stimuli. The gap between the two scales is of unknown magnitude. Since it would imply that the gap had some known magnitude if we were to begin the second scale with any other value, the new program once again starts with zero. The old program is incapable of obtaining the second scale from the input matrix. With the old program, we must run it twice in order to get both scales. The first run gives us five good values (a partial scale) and four meaningless ones. After deleting from the input matrix those data that produced the first partial scale, we must run the program again to get the final four values. When this is done, the resulting 4-by-4 matrix is of type M1 and yields a scale that begins with a negative number. However, the scales are equivalent.

Matrix (M4) gives us multiple disjoint scales. In this case, when the new program is used, there are 4 scale segments. The first consists of 3 stimuli (4, 1, and 5 in that order). These are separated by a large unmeasurable gap from the lone stimulus 6. This latter stimulus is separated by another large unmeasurable gap from stimuli 3 and 2, which are very close together but come in that order. Finally come the remaining 3 stimuli (9, 8, and 7), which are separated by a large unmeasurable gap from the previous 2 stimuli. As in the M3 case, the old program is incapable of doing more than assign scale values to the first 3 stimuli. However, if a new input matrix is prepared that eliminates the data for these 3 stimuli, the old program yields a second scale segment for stimuli 3 and 2, placing them in that order. When the data associated with these stimuli are also eliminated from the input matrix, the old program yields a third segment for stimuli 9, 8, 7, and 6 in that order. These results are summarized in Figure 11.

Step 4: Predicted Response Matrices

For both models (the scale produced by the old program and the one produced by the new program), we can use the scales to predict the responses that would be obtained in an assessment session with any number of observations. We do this by finding the difference between the scale values for any two stimuli and treating that difference as the unit normal standard

```
New Program          Old Program

1 ->  -0.35677       1 -> -0.361111
2 ->   0.206737      2 ->  0.210000
3 ->   0.160233      3 ->  0.161111
4 -> -0.409222       4 -> -0.413333
5 -> -0.341141       5 -> -0.345555
6 ->   0.398428      6 ->  0.403333
7 ->   0.255892      7 ->  0.258889
8 ->   0.0294193     8 ->  0.028889
```

Figure 8 — M1 scales

```
New Program          Old Program

 1 -> 0.000000        1 -> 0.000000
 2 -> 0.255073        2 -> 0.264620
 3 -> 0.343295        3 -> 0.350409
 4 -> 0.816674        4 -> 0.828804
 5 -> 0.343295        5 -> 0.350409
 6 -> 0.579156        6 -> 0.588304
 7 -> 0.115517        7 -> 0.117778
 8 -> 0.704121        8 -> 0.719804
 9 -> 1.4107          9 -> 1.434803
10 -> 0.115517       10 -> 0.117778
11 -> 0.226845       11 -> 0.230936
12 -> 0.829811       12 -> 0.847304
13 -> 1.05578        13 -> 1.070304
14 -> 1.93126        14 -> 1.964277
15 -> 0.710206       15 -> 0.725304
16 -> 1.10289        16 -> 1.122303
17 -> 0.343295       17 -> 0.350409
```

Figure 9 — M2 scales

```
New Program          Old Program

4 -> 0.000000        4 -> 0.000000
1 -> 0.480099        1 -> 0.484000
5 -> 0.628046        5 -> 0.638000
3 -> 1.20195         3 -> 1.213999
2 -> 1.24765         2 -> 1.261999

9 -> 0.000000        9 -> -0.547000
8 -> 0.532673        8 -> -0.017000
7 -> 0.730068        7 ->  0.210000
6 -> 0.909686        6 ->  0.355000
```

Figure 10 — M3 scales

```
New Program          Old Program

4 -> 0.000000        4 -> 0.000000
1 -> 0.481710        1 -> 0.488000
5 -> 0.628046        5 -> 0.638000

6 -> 0.000000        3 -> 0.000000
                     2 -> 0.060000
3 -> 0.000000
2 -> 0.047670        9 -> -0.55000
                     8 -> -0.02000
9 -> 0.000000        7 -> 0.210000
```

Figure 11— M4 scales

deviate for the differential preference for one of the stimuli over the other. A computer program already existed that implemented this process[9]. Given a scale, it gives a predicted response matrix that can be compared to the input matrix.

Figure 12 shows the predicted responses obtained from the two scales in Figure 8. We compare each of these to M1 (Figure 4) to obtain the mean squared error given by

$$MSE = \frac{\sum (m(i,j) - \hat{m}(i,j))^2}{n}, \tag{2}$$

where $m(i,j)$ is an entry from the original sum matrix (Figure 4), $\hat{m}(i,j)$ is an entry in the matrix of predicted responses (Figure 12), and n is the number of entries in the matrix for which the error is being computed. The results are shown in Table 4, not only for the M1 (complete, normal) case, but also for the remaining cases. We have omitted the predicted matrices for these other cases to conserve space. However, it is important to note here that when we get two disjoint scales, as in figure 10, they do not predict the entire response matrix. For example, in the M3 case, the first partial scale with 5 values, predicts 25 entries in a 5 x 5 subset of M3 and the second partial scale predicts 16 entries in a 4 x 4 subset of M3.

PREDICTED (New Program)

0	205	201	138	146	223	210	187	190
83	0	139	77	84	166	150	124	127
87	149	0	82	89	171	155	129	132
150	211	206	0	152	228	215	193	196
142	204	199	136	0	222	209	186	189
65	122	117	60	66	0	128	103	105
78	138	133	73	79	160	0	118	121
101	164	159	95	102	185	170	0	147
98	161	156	92	99	183	167	141	0

PREDICTED (Old Program)

0	206	201	138	146	224	211	188	191
82	0	138	77	83	166	150	123	127
87	150	0	81	88	172	155	129	132
150	211	207	0	152	228	216	193	196
142	205	200	136	0	223	209	186	189
64	122	116	60	65	0	127	102	105
77	138	133	72	79	161	0	118	121
100	165	159	95	102	186	170	0	147
97	161	156	92	99	183	167	141	0

Figure 12 – M1 responses predicted by scales from new program (left) and old program (right).

Thus, we can predict only 41 of the 81 values in M3. These are the values that produce the MSE shown in Table 4. The remaining 40 entries in M3 provided insufficient information to generate a scale and the scale contains insufficient information to predict the responses.

Step 5: Predicted Versus Actual

Using the mean squared error to compare the input matrix with each predicted matrix, we get the data in Table 4. We can see here that the scales from M2 and M4 produce predicted matrices that are very similar to the input matrix, but M3 gives larger MSEs and those for M1 are very large. The reason for this is simply that the scale is truly a best linear fit to the data. If the input matrix is very scattered, the scale will not be able to fit very well, and thus the scale cannot accurately predict the input data.

Table 4 – Mean Squared Error of the Old and New Programs Compared to the Input Matrix

Matrix	New Program Mean Squared Error	Old Program Mean Squared Error
M1	1254.5	1253.7
M2	0.085	0.065
M3	33.7	31.3
M4	0.22	0.32*

*If only 18 values are used for which the new program generated comparables, value is 0.22

Step 6: Old Model Versus New Model

In comparing the old model to the new model, let us use three factors: (1) the R-squared error, that is, the square of the Pearson correlation coefficient, (2) the mean squared error as presented previously in Table 4, and (3) the data comparing the elements of the predicted response matrices of the old and new programs.

R-Squared Comparison

First, by computing a regression line and R-squared as in Figure 13, we can see how closely the scale values of the old and new programs correlate. This was done for each of the scales in figures 8-11.

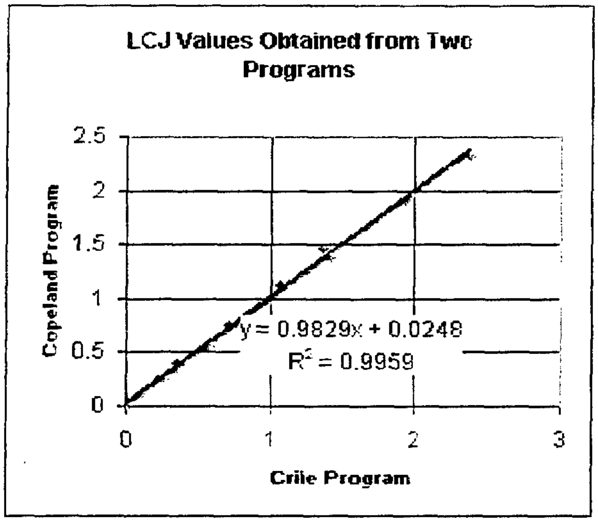


Figure 13 — Trend line for M2

The R-squared values are all very high: 1.0000, 0.9959, 1.0000, and 1.0000 respectively. In each case, Pearson's correlation coefficint is the square root of this value, with 1 indicating perfect correlation. This tells us that the two programs indeed are calculating equivalent scales, and the results they give are almost identical.

Mean Squared Error Comparison

In Table 4 above, it can be seen that the old program generally yields better (lower) mean square error. It performed better in 3 out of 4 cases.

Element Comparison

Comparing each element of each predicted response matrix to the original matrix to see which predicted response matrix had the greater error, we get the data in Table 5. Comparing the predicted matrices in this manner, we get a slightly different picture. Although the old program remains more accurate for M1, M2, and M3, we do see that for some elements, the new program was more accurate. Likewise, we get a different picture for M4 where there are the same number of entries for which each scale was more accurate.

Table 5 — Comparison of Individual Elements in Response Matrices

Matrix	New More Accurate	Old More Accurate
M1	12	18
M2	2	10
M3	10	16
M4	4	4

CONCLUSIONS AND RECOMMENDATIONS

While the new program does not yield exactly the same scale as the old one, these are "interval scales" and the same values are not required for equivalence. What is more important is the correlation between the scales produced by the two programs. After examining a number of cases that are believed to span the possibilities to be encountered, it has been

determined that there is an extremely high correlation between the scales produced by the two programs, with Pearson's correlation coefficient greater than 0.997 in all observed cases. However, comparing predicted responses from scales generated by the two programs appears to show that scales generated by the old program yield slightly better predictions.

The computational benefit expected from utilizing a continuous approximation for the area under the normal curve, in place of trapezoid-rule integration, did not materialize and may have actually added a systematic bias to the results. If such a bias exists, it is small.

In the multiple disjoint example, the old program produces no scale at all for 6 of the 9 stimuli. The improvement in the new program produces 4 scale segments, accounting for all of the stimuli. However, the results obtained are different than those obtained with the old program by the process described above. Which of these is the best representation of the data? At this point in time, the answer to this question is not clear. The new program considers all of the data available when placing stimulus 6 by itself between stimuli 5 and 3 and placing the stimuli in the order (4, 1, 5, 6, 3, 2, 9, 8, 7). Only the data relative to the 4 stimuli {6, 7, 8, 9} are considered when the old program places these in the order (9, 8, 7, 6) and provides a complete normal scale for them. If the new program is run on the same data, it produces an equivalent result. Further investigation is warranted in this matter.

With data such as that in Table 1 and Figure 1, Pearson's correlation coefficient is widely used and understood. However, there is no such metric that is known to the authors and implemented in software for describing how well an LCJ scale fits its underlying data. If there were one, it could be used to check the LCJ scales produced by the programs to determine whether the differences in the scales produced by the two programs were consistent and significant. In lieu of such a metric, in this paper we have used the mean squared error. It would be a good idea to investigate this issue further.

Finally, it is still necessary to test the part of this program that accepts the raw response data from the individual subjects and produces the sum matrix.

It would appear that the new program does calculate LCJ scales that are equivalent to those produced by the old program and incorporates some additional useful features. However, the algorithm of the old program seems as though it might be superior to that of the new one. Further investigation is warranted to determine whether there is indeed a bias in favor of the old program and, if there is, whether it is possible to change the algorithm of the new program to repair the defect. In the meanwhile, the new program could be used cautiously in place of the old one, being especially suspicious of multiple disjoint scales.

REFERENCES

[1] L. L. Thurstone, "A Law of Comparative Judgment," *Psychological Review*, 34:273-286 (1927).

[2] W. S. Torgerson, *Theory and Methods of Scaling*, John Wiley and Sons, New York (1958).

[3] W. S. Torgerson, *Theory and Methods of Scaling*, Krieger Publishing, Malabar, Florida (1985).

[4] A.C. Copeland, M.M. Trivedi, and G. Ravichandran, *Developing a Quantitative Basis for Synthesis, Analysis, and Assessment of Complex Camouflage Patterns*, U.S. Army Contract DAAK70-93-C-0037.

[5] E. L. Jacobs, "Modeling of Camouflage Screens Using Xpatch", *Proceedings of the Seventh Annual Ground Target Modeling and Validation Conference* (1996).

[6] J. R. McManamey, "Applying the Law of Comparative Judgement to Target Signature Evaluation", *RTO Meeting Proceedings 45, Search and Target Acquisition*, North Atlantic Treaty Organization RTO-MP-45 AC/323(SCI)TP/19 (2000).

[7] A. C. Copeland and M. M. Trivedi, Computational Models for Search Discrimination, Optical Engineering, 40:1885-1895 (2001).

[8] J.R. McManamey, "An Investigation of the Reliability of Search Statistics Based on Results from Paired Images," *Proceedings of the Thirteenth Annual Ground Target Modeling and Validation Conference*, Signature Research, Houghton (2002), In Press.

[9] J.R.McManamey, "LCJ2MAT.EXE" Computer program version 12/29/98, US Army, NVESD, Ft. Belvoir, VA (1998).

Non-Linear Mapping for Improvement of Display Comprehension of Low Resolution Images

John D. O'Connor and Jeffrey T. Olson
U. S Army CECOM RDEC
Night Vision and Electronic Sensors Directorate
Fort Belvoir, Virginia 22060-5806 USA

July 2, 2002

ABSTRACT

This paper describes the development and testing of Nonlinear Display Mapping (NDM), which is high-speed digital signal processing to quickly manipulate displayed images. The identification performance of observers using an automated algorithm and observers using NDM are compared. The observer utilizes NDM to improve image understanding. No algorithm or automatic method can optimize thermal displays for every environment, condition and target. Often, the manual controls are used when "auto mode" does not work well. The commonly available manual brightness and contrast controls are difficult to use and do not fully realize the potential of digital systems. Also, recent experiments showed that the auto mode might combine contrast shades such that targets or target features are hidden from the observer. Current display algorithms employ a wide variety of methods, including histogram equalization, local area processing, and region of interest processing. Non-linear Display Mapping (NDM) differs from these because it allows the user to manipulate the displayed intensity of different regions of the sensor output by real-time non-linear mapping to pixel values. The user can thus allocate or "tune" pixel intensities (gray shades) to output regions expected to contain targets. This avoids squandering the system's limited dynamic range on image features such as cold sky, clouds, trees or water. In other words, NDM enables the user to tune the sensor to the scene.

1. INTRODUCTION

The Night Vision Electronic Sensing Directorate Modeling Branch studied various sensor effects in the late 1990's through still-image degradation experiments. The effects of blur, sampling noise, spurious response and other variables on target identification were studied in experiments that used a standard set of 12-target 12-aspect high-resolution 12-bit thermal images as a baseline[1]. Once degraded, these images had to be optimized for 8-bit displays and it was found that no single technique could consistently deliver good contrast across the entire set of images while retaining critical target information. A novel method of non-linear mapping was devised that allowed the image processor to view the image histogram and determine non-linear bit allocations for display based on observable target features. Any target image could be quickly processed to achieve 50% contrast through this method.

Thermal sight users would benefit from such target contrast improvements if they could be achieved at a reasonable cost. Software was developed to simulate a simple non-linear sensor control. Over 900 high-, medium- and low-resolution thermal images with over 50 targets and 6 different backgrounds were processed for expert analysis. That analysis suggested that while non-linear mapping could improve target contrast for almost all studied conditions and targets, it would be most beneficial under conditions of moderate to high blur and moderate to low target resolvability. When blur is low and resolution

high, little improvement is possible because the observer has adequate target information to ID, regardless of target contrast. In conditions of very high blur and very low resolution, the target becomes an amorphous blob, and it is not possible to distinguish target features under any circumstances. Non-linear contrast enhancements are thus most beneficial between those extremes.

A low-resolution highly blurred target set with a predicted probability of ID under 30% tests the extreme limit of non-linear mapping's potential to improve image comprehension. The B-kit automatic algorithm emulator was chosen for comparison because it is scene based, known to produce relatively good target contrast under most conditions, and similar algorithms are incorporated into 2nd generation FLIR systems.

Night Vision Electronic Sensors Directorate researchers conducted a forced-choice vehicle ID experiment in which 20 trained subjects were shown low resolution images processed with a B-kit automatic algorithm emulator and with Non-linear Display Mapping (NDM) techniques. Subjects viewed 50 images processed with each method presented in a randomized order. The mean probabilities of identification for B-kit automatic and NDM were 25% and 32% respectively. Nineteen of 20 observers had higher scores with NDM, and the improvement range was 12% to 100%. The average improvement was 28%.

2. BACKGROUND AND PURPOSE

The 2nd generation Forward Looking Infrared (FLIR) imaging systems currently being developed and fielded have digital data output, whereas 1st generation FLIR's have analog output. Analog display controls generally allow the user to manipulate the brightness and contrast of the image, but such controls affect the entire data output signal, (generally 8 bit), and can introduce additional noise that obscures comprehension. While displays are often still limited to 8 bits, the output signal of the 2nd generation FLIR is twelve bits or more and, because it is digital, can be mapped to the display in numerous ways. Thus, regions of the 12-bit output signal are accentuated, or de-emphasized or eliminated when mapped to the 8-bit display.

Automatic algorithms and methods, such as region of interest and local area processing are useful image comprehension aids in some conditions, but the dynamic variability of thermal scenes and targets ultimately confounds these methods. The automatic method is not capable of prioritizing displayed gray shades based on subjective, situational or qualitative scene information. For example, significant gray shades may be allocated to a temperature gradient across the sky. When this occurs, image comprehension is lowered because gray shades that could have increased target internal contrast or other relevant information are allocated to the irrelevant sky temperature gradient.

The purpose of this Non-linear Display Mapping (NDM) research is to determine the level of image comprehension improvement gained by allowing the user to tune the sensor display to the scene. This is achieved by the user's active control of gray shade mapping to meet his comprehension or target acquisition needs. Figure 1 depicts

allocation of gray shades to expected target signal region. The user can thus employ subjective, situational and qualitative scene information to improve the displayed image. NDM does not increase perceived noise to achieve contrast improvement, as would be expected with analog gain and level controls. Thus, if sky temperature gradients, bodies of water, fires or hot pavement are interfering with scene comprehension, the user may "tune them out" by allocating a small or even single gray shade to those scene features. As the scene or task changes, the user can adapt the display to meet the current need. The user can also increase the gray shades allocated to objects of interest within the scene.

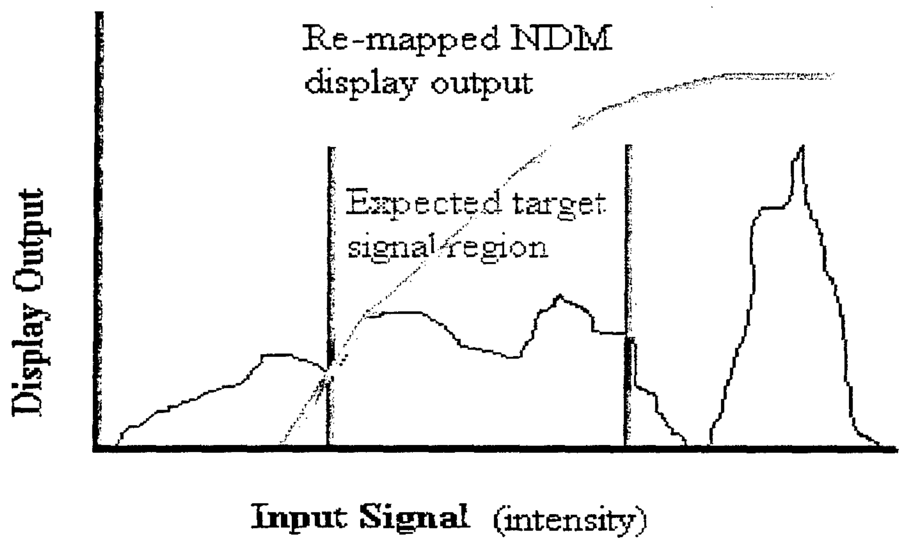


Figure 1. The NDM Concept

4. IMAGE SELECTION AND PROCESSING

Fifty frame-averaged 12-bit thermal images from an Agema 1000 longwave imager were selected for an ID experiment. Each image included one of nine tracked vehicle targets in the foreground, and mid-latitude dense forest in the background. The vehicles were: 2S3, BMP-1, M109A5, M113A2, M1IP, M2A2, T-62, T-72, and ZSU-23/4 (see Figure 2). Five to 7 of 8 different aspect orientations were selected for each vehicle (see Figure 3). This group of vehicle images represents a broad range of discrimination difficulty within the category of tracked vehicles for the identification task. This vehicle set also represents a good sample of the types of foreign and US tracked vehicles in current use. The images were collected at a range of 400m with a 20 x 13 degree field of view (FOV) and resolution of 590 x 401 pixels. The target in each image was centered, obviating the need for any search by the observer. The number of pixels on target was small, (from 160 to 400 pixels), occupying less than .2% of the total image area. It is important to note that the range and FOV selected represent a discrimination significantly beyond the expected 50% probability of ID (P(id)), where image information is highly constrained by blur (and thus lower target internal contrast) and resolution (pixels on target).

The 50 selected images were processed using a B-kit automatic algorithm emulator, developed by E-OIR measurements for the Night Vision Electronic Sensors Directorate (NVESD). The automatic algorithm emulator required no user input, and the display parameters were determined by the overall scene characteristics. Concurrently, the same 50 images were processed with NDM software developed by NVESD. The NDM software user viewed representative images and their histograms for each of the five collection environments from which images were drawn. The user then determined regions to be compressed or enhanced. Scene characteristics such as cold sky and hot ground were suppressed for the NDM processed images. Images from the same location and collection date were processed using the same NDM suppression and enhancement parameters to represent sensor output optimized for the scene, rather than for the target. This eliminated any bias based on user knowledge of the specific target in each image and its critical features or characteristics. The result of the two processing paths above yielded a total of 100 images (50 B-kit automatic and 50 NDM) that were presented to the subjects in a forced-choice vehicle ID computer perception test. This test software was initially developed by NVESD and has been used with a multitude of image sets to conduct over 20 different experiments over the last four years.

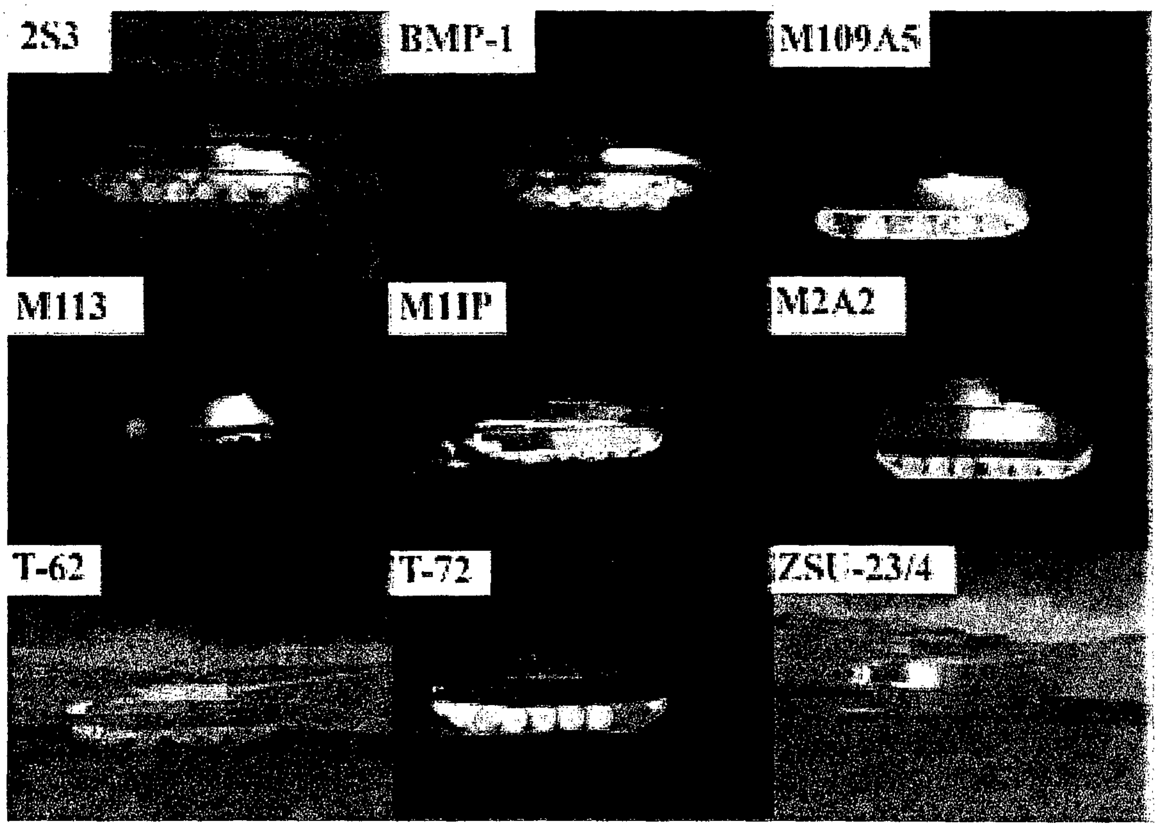


Figure 2. Vehicle targets types

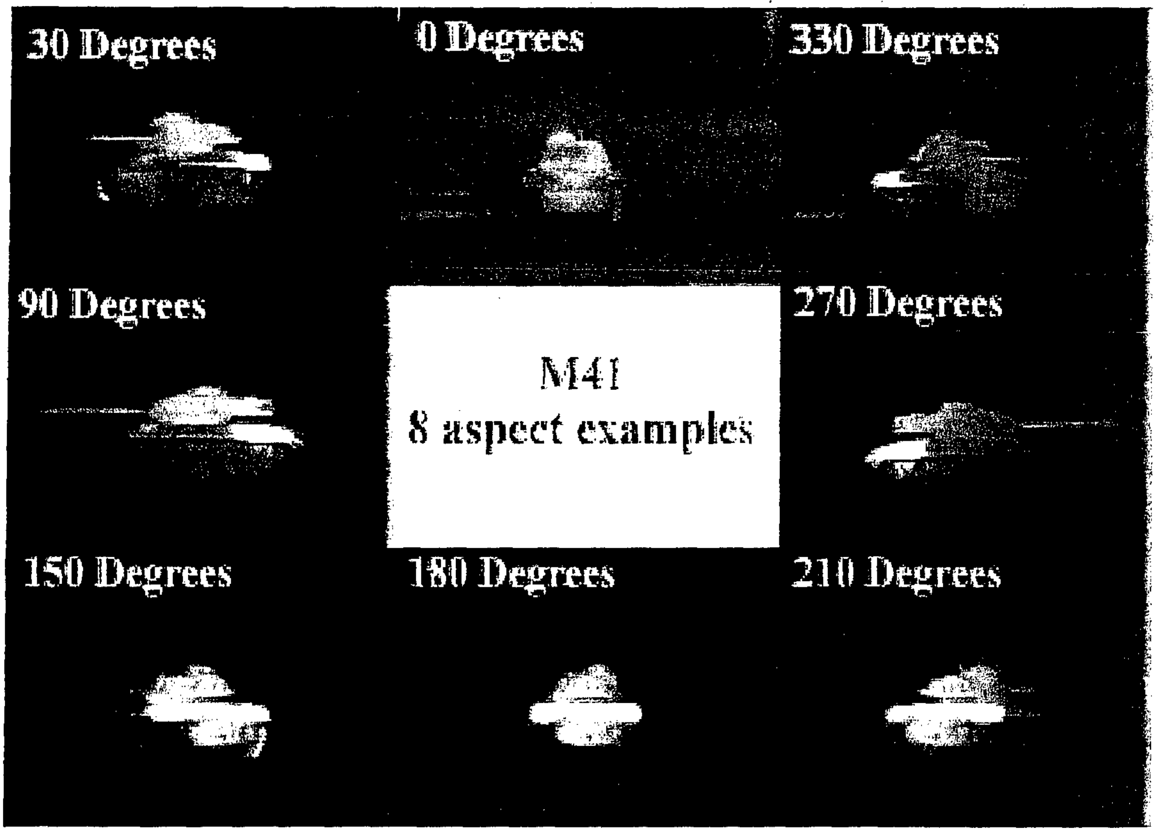


Figure 3. Vehicle image aspects

Figure 4 is a sample of the experimental imagery for each treatment, and A1 and A2 were generated from the same base image. It is easy to notice the relatively large number of gray shades allocated to the sky in A1, whereas in A2 the sky is reduced to a single gray shade. Note the greater internal target contrast of A2, and the separation of the target silhouette from the tree line. The greater target to background contrast and internal target contrast of A2 is made possible by the re-allocation of gray shades from cold sky temperatures, where there is little meaningful information, to expected target temperatures.

Images B1 and B2 in Figure 4 were generated from a base image different in location and environmental conditions from A1 and A2. The overcast sky in B1 and B2 is much closer to ambient temperature than in A1 and A2, where no clouds are evident. Note that the automatic algorithm emulator in B1 allocates many gray shades to the tree line but saturates the target, revealing little internal target detail. While the overall scene in B2 is lower contrast than B1, the target has good internal contrast, revealing the unsupported tracks, the location of the engine and exhaust, the turret crease, the folded radar dish, and other cues that distinguish this vehicle as a ZSU-23/4.

A1. M109A5, B-kit emulator A2. M109A5, NDM

B1. ZSU-23/4, B-kit emulator B2. ZSU-23/4, NDM

Figure 4. Low-resolution experimental image examples

5. EXPERIMENTAL METHOD

All test subjects were required to train to a 96% medium range P(id) standard (less than 1500 pixels on target) for the experimental set of nine vehicles before participating in this experiment. The US Army's *Recognition of Combat Vehicles*[2] (ROC-V) thermal signature training software was used to achieve this level of proficiency. Training thermal signatures is necessary for thermal ID experiments to reduce error introduced by guessing. For example, if a test subject is only 30% proficient in a set of 10 vehicles at close range (he can only positively ID 3 of 10), 70% or more of his ID calls for that set may be guesses. This introduces random error into the experimental data and reduces the reliability of results. Further, well-trained subjects can extract or deduce information from an image that less well-trained subjects cannot.

Subjects were individually administered the 100-image computer perception test discussed above upon completion of the required ROC-V training. Subjects were sequentially shown the images in randomized order and required to select the most likely vehicle identity based on what they could discern from the image. Images could not be

skipped or revisited. No time limit, per image or for the test overall, was imposed. Users received no feedback until the test's completion, and scores or any other aspects of the test were not discussed until all subjects had completed their tests. Subject test results were recorded as individual files for each individual. These files included subject identifiers, the image presentation order, and the scores, target calls and time between target calls for each image. It is important to note that the test imagery was collected separately from the ROC-V training imagery to avoid artifactual image identification, rather than the desired vehicle identification.

6. RESULTS

Twenty well-trained thermal sight users completed NVESD perception test. The mean scores for B-kit automatic and NDM images were 25.0 % (14% above chance) and 32.0% (21% above chance) respectively. Respective standard deviations were .78 and .74 (1.6% and 1.5%). The high and low scores for the B-kit treatment were 52% and 16%, with 70% of subjects scoring between 16% and 30%. The highest and lowest NDM treatment scores were 64% and 20%, with 70% of subjects scoring between 24% and 40%. Figure 5 illustrates each individual's ID performance with each treatment; with subjects ordered by B-kit treatment score. The error bars indicate that 19 of 20 subjects had significantly higher scores for the NDM treatment. NDM scores for those who improved were from 12% to 100% higher.

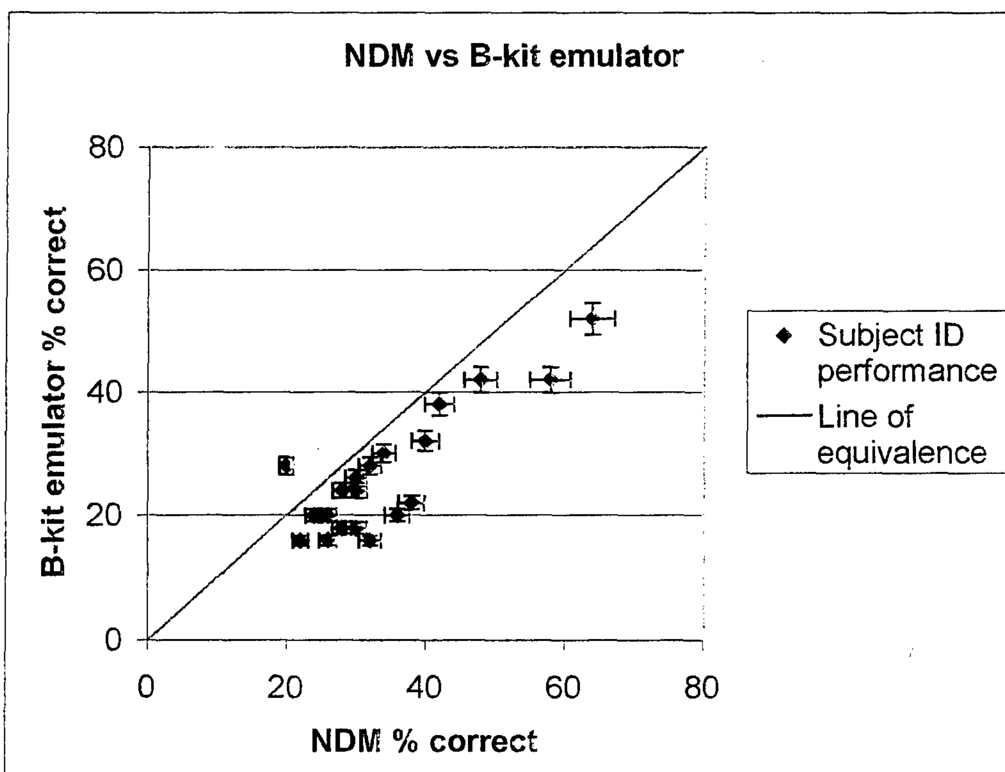


Figure 5. Low-resolution ID performance, NDM versus B-kit emulator

7. CONCLUSION

If the treatments were theoretically equal, the probability of 19 of 20 observers scoring higher with NDM would be 21 in 1,048,576. There is thus a better than 99.9% probability that NDM does improve comprehension with images having low resolution on target. Further experiments will measure the degree to which NDM can be expected to improve identification probabilities and reduce cycle criteria for N_{50}.[3]

REFERENCES:

1. Richard Vollmerhausen, Ronald Driggers, Barbara L. O'Kane. "Influence of sampling on target recognition and identification," *Optical Engineering*, **38** 5/763 May 1999
2. Recognition of Combat Vehicles (ROC-V) version 5E vehicle thermal signature trainer, Developed by NVESD, distributed by Project Manager Ground Combat Tactical Trainer (PM GCTT), sponsored by Joint Forces Identification Evaluation Team (JCIET) for US Department of Defense.
3. John Johnson, "Analysis of Image Forming Systems", Proc. of Image Intensifier Symposium, 1958, pp. 249-273.

ALLRED. KEN
ENGINEERING RESEARCH &
CONSULTING, INC.
555 SPARKMAN DR, STE. 1622
HUNTSVILLE AL 35898
kallred@rttc.army.mil

ANDERSON, LON
ARMY RESEARCH LAB
AMSRL-SL-EM
WHITE SANDS MISSILE RANGE NM 88002-5513
landerso@arl.army.mil

BALMA, ROBERT A
DEPT. OF NATIIONAL DEFENCE.
PROGRAMME MGMT. 3-7 101 COLONEL BY DRIVE
OTTAWA K1A 0K2 CANADA
Balma.RA@forces.ca

BENNETT. JOHN G
US ARMY TACOM
ATTN.: AMSTA-TR-R MS 263
WARREN MI 48397-5000
bennettj@tacom.army.mil

BLACKWELL, NEAL
US ARMY CECOM RESEARCH, DEV. &
ENG. CTR
FORT BELVOIR VA
Neal_E_Blackwell@belvoir.army.mil

BRYK, DARRYL C
US ARMY TARDEC
ATTN.: AMSTA-TR-R MS 263
WARREN MI 48397-5000
BrykD@tacom.army.mil

CLEMENT, DIETER
FGAN-FOM
Gutleuthausstrasse 1
Ettlingen D-76275
GERMANY
clement@fom.fgan.de

CRILE, JAMES
US ARMY TACOM
AMSTA-TR-R MS 263
WARREN MI 48397-5000

DENNEN, KEVIN
ERC INC.
555 SPARKMAN DRIVE
BUILDING 4500
HUNTSVILLE AL 35816
kdennen@rttc.redstone.army.mil

DERENIAK, EUSTACE
UNIVERSITY OF ARIZONA
1630 E. UNIVERSITY BLVD. OPTICAL
SCIENCES CTR
TUCSON AZ 85721
eustace@u.arizona.edu

EDWARDS, TIM
US ARMY CECOM
10221 BURBECK ROAD
FT. BELVOIR VA 22060
tedwards@nvl.army.mil

EVANS, ROGER
US ARMY TACOM
AMSTA-TR-R MS 263
WARREN MI 48397-5000
evansro@tacom.army.mil

FARRIER, JOHN
VERIDIAN ENGINEERING
5200 SPRINGFIELD STREET, STE. 200
DAYTON OH 45431
john.farrier@wpafb.af.mil

FRIEDMAN, MELVIN H.
NIGHT VISION & ELEC SENSOR DIR
10221 BURBECK RD, STE 430
AMSEL-RD-NV-VMD-SMT
FT. BELVOIR VA 22060
friedman@nvl.army.mil

GAMBORG, MARIUS
NORWEGIAN DEFENCE RESEARCH
ESTABLISHMENT
PO BOX 25
KJELLER NO-2027 NORWAY
marius.gamborg@ffi.no

GATLAND, DAVID
DSTL
AIR SYSTEMS A2 BUILDING, 1064
FARNBOROUGH HAMPSHIRE
GU14 0LX UK
DGGatland@mail.dstl.gov.uk

GERHARDT, ALICE
THERMOANALYTICS, INC
23440 AIRPARK BLVD.
CALUMET MI 49913
amg@thermoanalytics.com

GERHART, GRANT
US ARMY TARDEC
ATTN: AMSTA-TR-R MS 263
WARREN MI 48397-5000
gerhartg @tacom.army.mil

GOLDSWORTHY, B. JANE
DSTL
AIR SYSTEMS A2 BUILDING, 1064
FARNBOROUGH HAMPSHIRE
GU14 0LX UK
jbgoldsworthy@dstl.gov.uk

GONDA, TERESA
US ARMY TACOM
ATTN: AMSTA-TR-R MS 263
WARREN MI 48397-5000
gondat@tacom.army.mil

HAMMON, RICKY
US ARMY AMCOM
AMSAM-RD-MG-IP
REDSTONE ARSENAL AL 35898
ricky.hammon@rdec.redstone.army.mil

HUTCHINSON, HAROLD R
US ARMY TACOM
Attn: AMSTA-CM-KSF MS 105
WARREN MI 48397-5000
hutchinh@tacom.army.mil

JACKSON, WILLIAM
US ARMY TARDEC
AMSTA-TR-R, 6501 E. 11 MILE ROAD
MS 263
WARREN MI 48397-5000
JacksonW@tacom.army.mil

JOHNSON, KEITH
THERMOANALYTICS, INC.
23440 AIRPARK BLVD.
CALUMET MI 49913
krj@ThermoAnalytics.com

JONES, JACK
US ARMY TACOM
Attn: AMSTA-TR-R
WARREN MI 48397-5000
jack@qix.net

KANG, ROBIN
SAIC
4001 FAIRFAX DRIVE
ARLINGTON VA 22203
rkang@nvl.army.mil

KARLSEN, ROBERT
US ARMY TACOM
ATTN: AMSTA-TR-R MS 263
WARREN MI 48397-5000
karlsenr@tacom.army.mil

KIMBEL, CARRIE
DYNETICS, INC.
PO BOX 5500
HUNTSVILLE AL 35814-5500
carrie.kimbel@dynetics.com

KOENIG, GEORGE
US ARMY CRREL
72 LYME RD
HANOVER NY 03755-1290
George.G.Koenig@ERDC.usace.army.mil

LABATAILLE, ROGER J.
US ARMY TACOM
ATTN: SFAE-ASM-SS
WARREN MI 48397-5000

LANE, LORI
VERIDIAN ENGINEERING
5200 SPRINGFIELD STREET. STE. 200
DAYTON OH 45431
lori.lane@wpafb.af.mil

LEAP, DARREN
GENERAL DYNAMICS AMPHIBIOUS SYSTEMS
991 ANNAPOLIS WAY
WOODBRIDGE VA 22191

LE, CALVIN
US ARMY RESEARCH LAB
2800 POWDER MILL ROAD
ADELPHI MD 20783
cle@arl.army.mil

LIEBOWITZ, DAVID
US ARMY MATERIAL SYS. ANAL.
ATTN: AMXSY-CA
392 HOPKINS ROAD
APG MD 21005
Liebgwd@amsaa.army.mil

LONDINO, JOSEPH
KILGORE FLARE
155 KILGORE DRIVE
TOONE TN 38381

LOYD, JODY S.
SIMULATION TECHNOLOGIES, INC.
PO BOX 7009
HUNTSVILLE AL 35807
Jody.Loyd@rdec.redstone.army.mil

MANTEY, ROBERT
US ARMY TACOM
6501 E. 11 Mile Rd., AMSTA-TR-R
MS 263
WARREN MI 48397-5000
ManteyR@tacom.army.mil

MCMANAMEY, JAMES R
NIGHT VISION & ELEC SENSOR DIR
10221 BURBECK RD. STE 430
AMSEL-RD-NV-ST-VMS
FT. BELVOIR VA 22060-5806
jmcmanam@nvl.army.mil

MEANS, JAMES
TEMEKU TECHNOLOGIES INC.
PO BOX 722
LUTZ FL 33548-0722
jlmeans@earthlink.net

MEITZLER, THOMAS
US ARMY TACOM
6501 E. 11 MILE ROAD,
ATTN: AMSTA-TR-R MS 263
WARREN MI 48397-5000
meitzlet@tacom.army.mil

MILLER, BRIAN
NVESD
10221 BURBECK ROAD
FT. BELVOIR VA 22060
bsmiller@nvl.army.mil

NEWLIN, DR. D. DAVID
GDLS
38500 MOUND ROAD
STERLING HEIGHTS MI 48310
newlind@gdls.com

O'CONNOR, JOHN
NVESD
10221 BURBECK RD.
FT. BELVOIR VA 22060

PASSMORE, RONALD
US ARMY AMCOM
ATTN: AMSAM-RD-MG-IR
REDSTONE ARSENAL AL 35898
ronald.passmore@rdec.redstone.army.mil

PEREZ ,JEFFREY
US ARMY TARDEC
AMSTA-TR-R MS 263
WARREN MI 48397-5000
perezj@tacom.army.mil

PETIT, GEORGE
RAYTHEON COMPANY-ARL MSRC
939-I BEARDS HILL RD, PMB 191
ABERDEEN MD 21001
gpetit@arl.army.mil

PICKARD, WARREN
AMTEC CORPORATION
500 WYNN DRIVE. SUITE 314
HUNTSVILLE AL 35816
wpickard@amtec-corp.com

RAPANOTTI, JOHN
DRDC-VALCARTIER
2459 BOUL. PIE-XI BLVD. NORTH
VAL-BELAIR QC G3J 1X5
CANADA
john.rapanotti@drev.dnd.ca

REED, JACK
US ARMY TACOM, TARDEC
AMSTA-TR-R, 6501 E. 11 MILE ROAD
MS 263
WARREN MI 48397-5000
reedj@tacom.army.mil

ROUSE, WILLIAM
US ARMY EDGEWOOD CHEM AND BIO CTR
ATTN.: AMSSB-RRT-TD E3330/155
ABERDEEN PROV. GRD MD 21010
William.Rouse@sbccom.apgea.army.mil

RYZYI, MARK G.
US ARMY TACOM, PM ABRAMS SYS
ATTN: SFAE-GCSS-AB-S
WARREN MI 48397
ryzyim@tacom.army.mil

SANDERS, JOHNNIE
LOCKHEED MARTIN
5600 SAND LAKE ROAD MP-951
ORLANDO FL 32819
johnnie.j.sanders@lmco.com

SAUTER, DAVID P
ARL, BATTLEFIELD ENV DIR.
ATTN: AMSRL-CI-EW
WHITE SANDS MISSILE NM 88002-5501
dsauter@arl.army.mil

SHAMAI, SHLOMO
ELOP
PO Box 1165
REHOVOT 76110
ISRAEL
shamais@bezeqint.net

SHMUELI, GAL
ELECTRO-OPTICS RESEARCH & DEV.
EORD LTD.
TECHNION CITY HAIFA 32000
ISRAEL
g95_2000@yahoo.com

SLOSS, DAVID
US ARMY TACOM
ATTN: AMSTA-TR-N MS 289
WARREN MI 48397
slossd@tacom.army.mil

SPURGEON, WILLIAM
US ARMY RESEARCH LAB
AMSRL-WM-MD
APG MD 21005-5069
wspurge@arl.army.mil

THOMAS, DAVID
US ARMY TACOM
ATTN: AMSTA-TR-R MS 263
WARREN MI 48397-5000
thomadav@tacom.army.mil

TRUAX, JAMES
THERMOANALYTICS, INC.
23440 AIRPARK BLVD.
CALUMET MI 49913
jdt@thermoanalytics.com

VANDEREEDT, ARTHUR
MILTEC CORPORATION
678 DISCOVERY DRIVE
HUNTSVILLE AL 35806
vandereedtaj@navair.navy.mil

VALISETTY. RAMA
US ARMY RESEARCH LAB
AMSRL-OP-AP-1 #MAILROOM
APG MD 21005
Rama@arl.army.mil

VOLLMERHAUSEN, RICHARD
NIGHT VISION & ELEC SENSOR DIR
10221 BURBECK RD, STE 430
AMSEL-RD-NV-AST
FT. BELVOIR VA 22060-5677
richard.vollmerhausen@nvl.army.mil

WEBER. BRUCE
US ARMY RESEARCH LAB
2800 POWDERMILL RD
AMSRL-SE-SE
ADELPHI MD 20783
bweber@arl.mil

WILLIAMS, RANDALL
US ARMY ERDC
3909 HALLS FERRY ROAD
VICKSBURG MS 39180
williar@wes.army.mil

WITUS, GARY
TURING ASSOCIATES INC
1392 HONEY RUN DRIVE
ANN ARBOR MI 48103
witusg@umich.edu